(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 January 2004 (22.01.2004)

PCT

(10) International Publication Number
WO 2004/007773 A1

(51) International Patent Classification[7]: C12Q 1/68, C12P 19/34, C07H 21/02, 21/04, 19/00

(21) International Application Number: PCT/US2003/021818

(22) International Filing Date: 11 July 2003 (11.07.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/194,882 12 July 2002 (12.07.2002) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 10/194,882 (CIP)
Filed on 12 July 2002 (12.07.2002)

(71) Applicant *(for all designated States except US)*: **THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK** [US/US]; West 116th Street and Broadway, New York, NY 10027 (US).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: **JU, Jingyue** [US/US]; 267 Marietta Street, Englewood Cliffs, NJ 07632 (US).

(74) Agent: **WHITE, John, P.**; Cooper & Dunham LLP, 1185 Avenue of the Americas, New York, NY 10036 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: MULTIPLEX GENOTYPING USING SOLID PHASE CAPTURABLE DIDEOXYNUCLEOTIDES AND MASS SPECTROMETRY

(57) Abstract: This invention provides methods for detecting single nucleotide polymorphisms and multiplex genotyping using dideoxynucleotides and mass spectrometry.

WO 2004/007773 A1

Applicants: Jingyue Ju
Serial No.: 10/521,206
Filed: November 9, 2006
Exhibit 2

MULTIPLEX GENOTYPING USING SOLID PHASE CAPTURABLE DIDEOXYNUCLEOTIDES AND MASS SPECTROMETRY

This application is a continuation-in-part and claims priority of U.S. Serial No. 10/194,882, filed July 12, 2002, the contents of which are hereby incorporated by reference into this application.

Background Of The Invention

Throughout this application, various publications are referenced in parentheses. Citations for these references may be found at the end of the specification immediately preceding the claims. The disclosures of these publications in their entireties are hereby incorporated by reference into this application to more fully describe the state of the art to which this invention pertains.

Single nucleotide polymorphisms (SNPs), the most common genetic variations in the human genome, are important markers for identifying disease genes and for pharmacogenetic studies (1, 2). SNPs appear in the human genome with an average density of once every 1000-base pairs (3). To perform large-scale SNP genotyping, a rapid, precise and cost-effective method is required. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) (4) allows rapid and accurate sample measurements (5-7) and has been used in a variety of SNP detection methods including hybridization (8-10), invasive cleavage (11, 12) and single base extension (SBE) (5, 13-17). SBE is widely used for multiplex SNP analysis. In this method, primers

designed to anneal immediately adjacent to a polymorphic site are extended by a single dideoxynucleotide that is complementary to the nucleotide at the variable site. By measuring the mass of the resulting extension product, a particular SNP can be identified. Current SBE methods to perform multiplex SNP analysis using MS require unambiguous simultaneous detection of a library of primers and their extension products. However, limitations in resolution and sensitivity of MALDI-TOF MS for longer DNA molecules make it difficult to simultaneously measure DNA fragments over a large mass range (6). The requirement to measure both primers and their extension products in this range limits the scope of multiplexing.

A high fidelity DNA sequencing method has been developed which uses solid phase capturable biotinylated dideoxynucleotides (biotin-ddNTPs) by detection with fluorescence (18) or mass spectrometry (19), eliminating false terminations and excess primers. Combinatorial fluorescence energy transfer tags and biotin-ddNTPs have also been used to detect SNPs (20).

False stops or terminations occur when a deoxynucleotide rather than a dideoxynucleotide terminates a se+quencing fragment. It has been shown that false stops and primers which have dimerized can produce peaks in the mass spectra that can mask the actual results preventing accurate base identification (21).

The present application discloses an approach using solid phase capturable biotin-ddNTPs in SBE for multiplex genotyping by MALDI-TOF MS. In this method primers that

have different molecular weights and that are specific to the polymorphic sites in the DNA template are extended with biotin-ddNTPs by DNA polymerase to generate 3'-biotinylated DNA extension products. The 3'-biotinylated DNAs are then captured by streptavidin-coated magnetic beads, while the unextended primers and other components in the reaction are washed away. The pure DNA extension products are subsequently released from the magnetic beads, for example by denaturing the biotin-streptavidin interaction with formamide, and analyzed with MALDI-TOF MS. The nucleotide at the polymorphic site is identified by analyzing the mass difference between the primer extension product and an internal mass standard added to the purified DNA products. Since the primer extension products are isolated prior to MS analysis, the resulting mass spectrum is free of non-extended primer peaks and their associated dimers, which increases the accuracy and scope of multiplexing in SNP analysis. The solid phase purification system also facilitates desalting of the captured oligonucleotides. Desalting is critical in sample preparation for MALDI-TOF MS measurement since alkaline and alkaline earth salts can form adducts with DNA fragments that interfere with accurate peak detection (21).

Summary Of The Invention

This invention is directed to a method for determining the identity of a nucleotide present at a predetermined site in a DNA whose sequence immediately 3′ of such predetermined site is known which comprises:

(a) treating the DNA with an oligonucleotide primer whose sequence is complementary to such known sequence so that the oligonucleotide primer hybridizes to the DNA and forms a complex in which the 3′ end of the oligonucleotide primer is located immediately adjacent to the predetermined site in the DNA;

(b) simultaneously contacting the complex from step (a) with four different labeled dideoxynucleotides, in the presence of a polymerase under conditions permitting a labeled dideoxynucleotide to be added to the 3′ end of the primer so as to generate a labeled single base extended primer, wherein each of the four different labeled dideoxynucleotides (i) is complementary to one of the four nucleotides present in the DNA and (ii) has a molecular weight which can be distinguished from the molecular weight of the other three labeled dideoxynucleotides using mass spectrometry; and

(c) determining the difference in molecular weight between the labeled single base extended primer and the oligonucleotide primer so as to identify the dideoxynucleotide incorporated into the single base extended primer and thereby determine the identity of the

nucleotide present at the predetermined site in the DNA.

In one embodiment, the method further comprises after step (b) the steps of:

(i) contacting the labeled single base extended primer with a surface coated with a compound that specifically interacts with a chemical moiety attached to the dideoxynucleotide by a linker so as to thereby capture the extended primer on the surface; and

(ii) treating the labeled single base extended primer so as to release it from the surface.

In one embodiment, the method further comprises after step (i) the step of treating the surface to remove primers that have not been extended by a labeled dideoxynucleotide.

Brief Description Of The Figures

Figure 1: Scheme of single base extension for multiplex SNP analysis using biotin-ddNTPs and MALDI-TOF MS. Primers that anneal immediately next to the polymorphic sites in the DNA template are extended by DNA polymerase of a biotin-ddNTP in a sequence-specific manner. After solid phase capture and isolation of the 3'-biotinylated DNA extension fragments, MALDI-TOF MS was used to analyze these DNA products to yield a mass spectrum. From the relative mass of each extended primer, compared to the mass of an internal standard, the nucleotide at the polymorphic site is identified.

Figure 2. Multiplex SNP genotyping mass spectra generated using biotin-ddNTPs. Inset is a magnified view of heterozygote peaks. Masses of the extension product in reference to the internal mass standard were listed on each single base extension peak. The mass values in parenthesis indicate the mass difference between the extension products and the corresponding primers. (A) Detection of six nucleotide variations from synthetic DNA templates mimicking mutations in the *p53* gene. Four homozygous (T, G, C and C) and one heterozygous (C/A) genotypes were detected. (B) Detection of two heterozygotes (A/G and C/G) in the human *HFE* gene.

Figure 3: Structure of four mass tagged biotinylated ddNTPs. Any of the four ddNTPs (ddATP, ddCTP, ddGTP, ddTTP) can be used with any of the illustrated linkers.

Figure 4: Synthesis scheme for mass tag linkers. For illustrative purposes, the linkers are labeled to

correspond to the specific ddNTP with which they are shown coupled in Figures 3, 5, 7, 8 and 9. However, any of the three linkers can be used with any ddNTP. (i) $(CF_3CO)_2O$; (ii) Disuccinimidylcarbonate/ diisopropylethylamine; (iii) Propargyl amine.

Figure 5: The synthesis of ddATP-Linker-II-11-Biotin. (i) Linker II, tetrakis(triphenylphosphine) palladium(0); (ii) $POCl_3$, Bn_4N^+ pyrophosphate; (iii) NH_4OH; (iv) Sulfo-NHS-LC-Biotin.

Figure 6: DNA products are purified by a streptavidin coated porous silica surface. Only the biotinylated fragments are captured. These fragments are then cleaved by light irradiation (*hv*) to release the captured fragments, leaving the biotin moiety still bound to the streptavidin.

Figure 7: Mechanism for the cleavage of photocleavable linkers.

Figure 8: The structures of ddNTPs linked to photocleavable (PC) biotin. Any of the four ddNTPs (ddATP, ddCTP, ddGTP, ddTTP) can be used with any of the shown linkers.

Figure 9: The synthesis of ddATP-Linker-II-PC-Biotin. PC = photocleavable.

Figure 10: Schematic for capturing a DNA fragment terminated with a dideoxynucleoside monophosphate on a surface. The dideoxynucleoside monophosphate (ddNMP) which is on the 3′ end of the DNA fragment is attached

via a linker to a chemical moiety "X" which interacts with a compound "Y" on the surface to capture the DNA fragment terminated with the ddNMP. The DNA fragment can be freed from the surface either by disrupting the interaction between chemical moiety X and compound Y (lower scheme) or by cleaving the linker (upper scheme).

Figures 11A-11C: Schematic of a high throughput channel based purification system. Sample solutions can be pushed back and forth between the two plates through glass capillaries and the streptavidin coated channels in the chip. The whole chip can be irradiated to cleave the samples after immobilization.

Figure 12: The synthesis of streptavidin coated porous surface.

Figures 13A-13C: Simultaneous detection of nucleotide variations in 30 codons of the *p53* tumor suppressor gene by MALDI-TOF MS using solid phase capturable biotinylated dideoxynucleotide. Each peak represents a different polymorphism labeled with its nucleotide identity and absolute mass value. The value in parentheses, denoting the mass difference between each DNA extension product and its corresponding primer, is used to determine the nucleotide identity. (**A**) A mass spectrum from a Wilms' tumor sample showing 30 wild type *p53* sequences. (**B**) A mass spectrum from a head and neck tumor (primary tumor biopsy) containing a heterozygous genotype G/T (4684/4734 Da) (boxed) in codon 157, corresponding to the wild type and mutant alleles, respectively. (**C**) A mass spectrum from a colorectal tumor cell line (HT-29) containing a homozygous G to A mutation (boxed) in codon 273 of the

p53 gene. The colorectal tumor cell line (SW-480) contained the identical G to A mutation in codon 273.

Figures 14A-14B: (**A**) A mass spectrum from a head and neck tumor sample showing 30 wild type sequences of the *p53* gene. (**B**) A mass spectrum from a head and neck tumor cell line (SCC-4) containing a homozygous C (5881 Da) to T (5970 Da) mutation (boxed) in codon 151 of the *p53* gene. Both spectra were produced using the primers shown in Table 3 with primer 16 replaced by primer 5'-TGTGGGTTGATTCCACA-3' for detecting the variation in codon 151 (C/TCC).

Detailed Description Of The Invention

The following definitions are presented as an aid in understanding this invention.

The standard abbreviations for nucleotide bases are used as follows: adenine (A), cytosine (C), guanine (G), thymine (T), and uracil (U).

A nucleotide analogue refers to a chemical compound that is structurally and functionally similar to the nucleotide, i.e. the nucleotide analogue can be recognized by polymerase as a substrate. That is, for example, a nucleotide analogue comprising adenine or an analogue of adenine should form hydrogen bonds with thymine, a nucleotide analogue comprising C or an analogue of C should form hydrogen bonds with G, a nucleotide analogue comprising G or an analogue of G should form hydrogen bonds with C, and a nucleotide analogue comprising T or an analogue of T should form hydrogen bonds with A, in a double helix format.

This invention is directed to a method for determining the identity of a nucleotide present at a predetermined site in a DNA whose sequence immediately 3′ of such predetermined site is known which comprises:

(a) treating the DNA with an oligonucleotide primer whose sequence is complementary to such known sequence so that the oligonucleotide primer hybridizes to the DNA and forms a complex in which the 3′ end of the oligonucleotide primer is located immediately adjacent to the predetermined site in the DNA;

(b) simultaneously contacting the complex from step (a) with four different labeled dideoxynucleotides, in the presence of a polymerase under conditions permitting a labeled dideoxynucleotide to be added to the 3' end of the primer so as to generate a labeled single base extended primer, wherein each of the four different labeled dideoxynucleotides (i) is complementary to one of the four nucleotides present in the DNA and (ii) has a molecular weight which can be distinguished from the molecular weight of the other three labeled dideoxynucleotides using mass spectrometry; and

(c) determining the difference in molecular weight between the labeled single base extended primer and the oligonucleotide primer so as to identify the dideoxynucleotide incorporated into the single base extended primer and thereby determine the identity of the nucleotide present at the predetermined site in the DNA.

In one embodiment, each of the four labeled dideoxynucleotides comprises a chemical moiety attached to the dideoxynucleotide by a different linker which has a molecular weight different from that of each other linker.

In one embodiment, the method further comprises after step (b) the steps of:

(i) contacting the labeled single base extended primer with a surface coated with a compound

that specifically interacts with a chemical moiety attached to the dideoxynucleotide by a linker so as to thereby capture the extended primer on the surface; and

(ii) treating the labeled single base extended primer so as to release it from the surface.

In a further embodiment, the method comprises after step (i) the step of treating the surface to remove primers that have not been extended by a labeled dideoxynucleotide and any non-captured component.

In one embodiment of the method step (c) comprises determining the difference in mass between the labeled single base extended primer and an internal mass calibration standard added to the extended primer. In one embodiment, the internal mass standard is 5'-TTTTTCTTTTTCT-3' (SEQ ID NO: 5) (MW = 3855 Da).

In one embodiment, the chemical moiety is attached via a different linker to different dideoxynucleotides. In one embodiment, the different linkers increase mass separation between different labeled single base extended primers and thereby increase mass spectrometry resolution.

In one embodiment, the dideoxynucleotide is selected from the group consisting of 2',3'-dideoxyadenosine 5'-triphosphate (ddATP), 2',3'-dideoxyguanosine 5'-triphosphate (ddGTP), 2',3'-dideoxycytidine 5'-triphosphate (ddCTP), and 2',3'-dideoxythymidine 5'-triphosphate (ddTTP).

In different embodiments of the methods described herein, the interaction between the chemical moiety attached to the dideoxynucleotide by the linker and the compound on the surface comprises a biotin-streptavidin interaction, a phenylboronic acid-salicylhydroxamic acid interaction, or an antigen-antibody interaction.

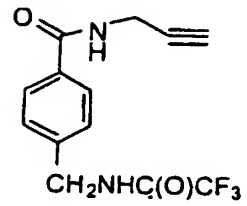
In one embodiment, the step of releasing the labeled single base extended primer from the surface comprises disrupting the interaction between the chemical moiety attached by the linker to the dideoxynucleotide and the compound on the surface. In different embodiments, the interaction is disrupted by a means selected from the group consisting of one or more of a physical means, a chemical means, a physical chemical means, heat, and light. In one embodiment, the interaction is disrupted by light. In one embodiment, the interaction is disrupted by ultraviolet light. In different embodiments, the interaction is disrupted by ammonium hydroxide, formamide, or a change in pH (-log H^+ concentration).

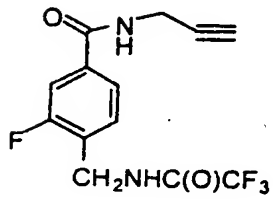
In different embodiments, the linker can comprise a chain structure, or a structure comprising one or more rings, or a structure comprising a chain and one or more rings. In different embodiments, the dideoxynucleotide comprises a cytosine or a thymine with a 5-position, or an adenine or a guanine with a 7-position, and the linker is attached to the dideoxynucleotide at the 5-position of cytosine or thymine or at the 7-position of adenine or guanine.

In different embodiments, the step of releasing the labeled single base extended primer from the surface comprises cleaving the linker between the chemical moiety and the dideoxynucleotide. In different embodiments, the linker is cleaved by a means selected from the group consisting of one or more of a physical means, a chemical means, a physical chemical means, heat, and light. In one embodiment, the linker is cleaved by light. In one embodiment, the linker is cleaved by ultraviolet light. In different embodiments, the linker is cleaved by ammonium hydroxide, formamide, or a change in pH (-log H^+ concentration).

In one embodiment, the linker comprises a derivative of 4-aminomethyl benzoic acid. In one embodiment, the linker comprises a 2-nitrobenzyl group or a derivative of a 2-nitrobenzyl group. In one embodiment, the linker comprises one or more fluorine atoms.

In one embodiment, the linker is selected from the group consisting of:





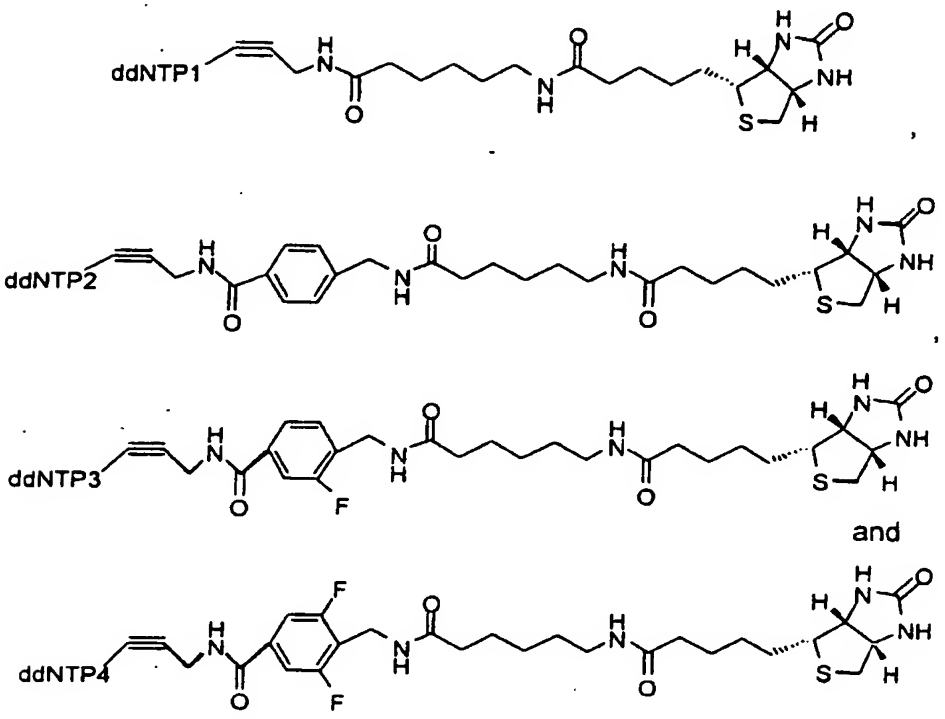
and

In one embodiment, a plurality of different linkers is used to increase mass separation between different labeled single base extended primers and thereby increase mass spectrometry resolution.

In one embodiment, the chemical moiety comprises biotin, the labeled dideoxynucleotide is a biotinylated dideoxynucleotide, the labeled single base extended

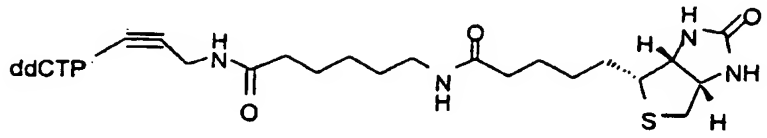
primer is a biotinylated single base extended primer, and the surface is a streptavidin-coated solid surface. In one embodiment, the biotinylated dideoxynucleotide is selected from the group consisting of ddATP-11-biotin, ddCTP-11-biotin, ddGTP-11-biotin, and ddTTP-16-biotin.

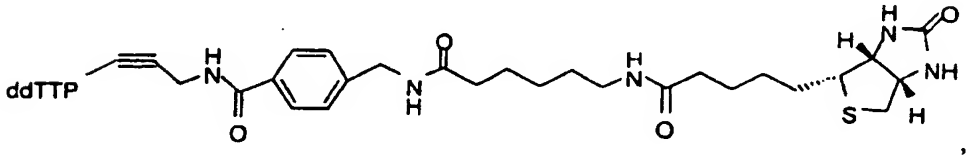
In one embodiment, the biotinylated dideoxynucleotide is selected from the group consisting of:

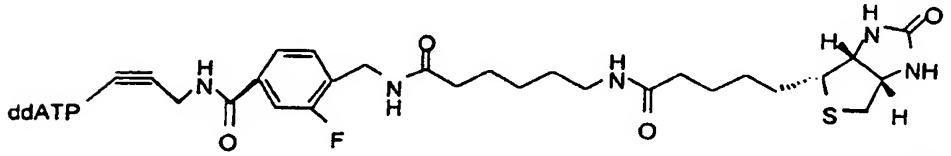


wherein ddNTP1, ddNTP2, ddNTP3, and ddNTP4 represent four different dideoxynucleotides, or their analogues.

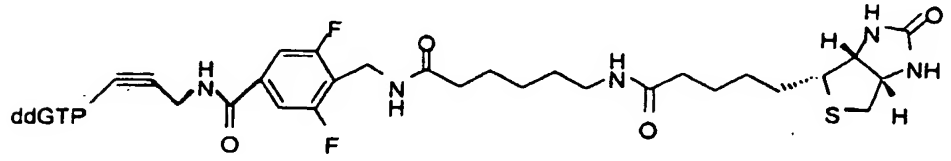
In one embodiment, the biotinylated dideoxynucleotide is selected from the group consisting of:



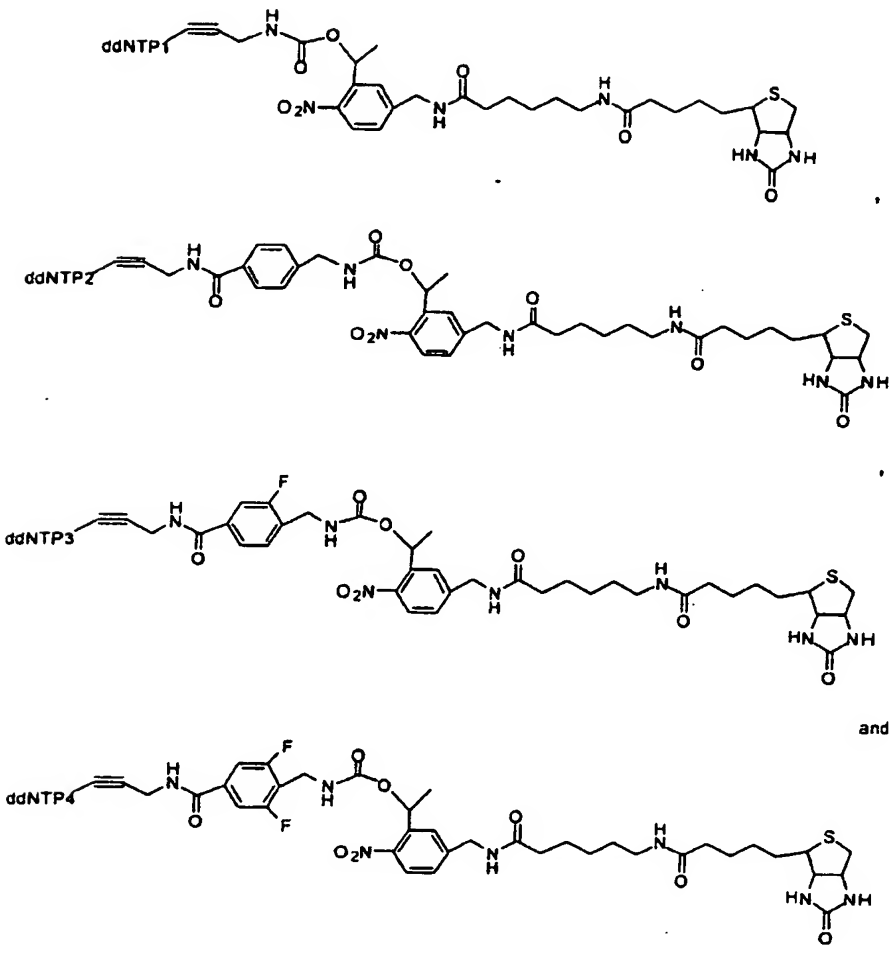




and

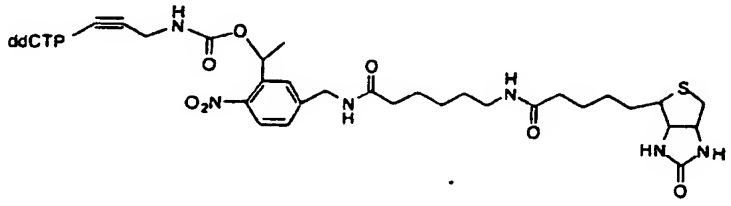


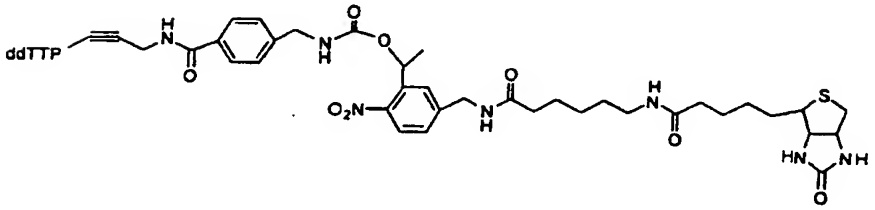
In one embodiment, the biotinylated dideoxynucleotide is selected from the group consisting of:

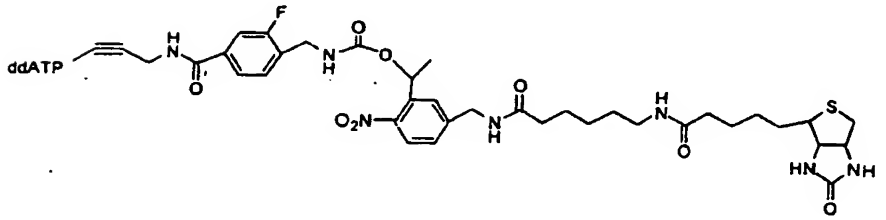


wherein ddNTP1, ddNTP2, ddNTP3, and ddNTP4 represent four different dideoxynucleotides or their analogues.

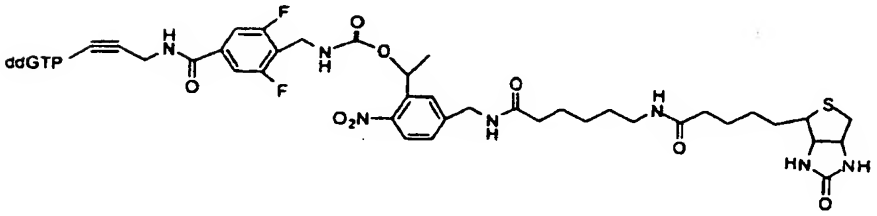
In one embodiment, the biotinylated dideoxynucleotide is selected from the group consisting of:







and



In one embodiment, the streptavidin-coated solid surface is a streptavidin-coated magnetic bead or a streptavidin-coated silica glass.

In one embodiment of the method, steps (a) and (b) are performed in a single container or in a plurality of connected containers.

The invention provides methods for determining the identity of nucleotides present at a plurality of predetermined sites, which comprises carrying out any of the methods disclosed herein using a plurality of different primers each having a molecular weight different from that of each other primer, wherein a different primer hybridizes adjacent to a different predetermined site. In one embodiment, different linkers each having a molecular weight different from that of each other linker are attached to the different dideoxynucleotides to increase mass separation between different labeled single base extended primers and thereby increase mass spectrometry resolution.

In one embodiment, the mass spectrometry is matrix-assisted laser desorption/ionization time-of-flight mass spectrometry.

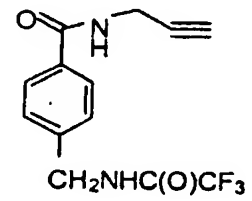
Linkers are provided for attaching a chemical moiety to a dideoxynucleotide, wherein the linker comprises a derivative of 4-aminomethyl benzoic acid.

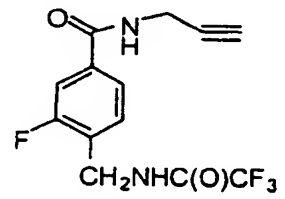
In one embodiment, the dideoxynucleotide is selected from the group consisting of 2',3'-dideoxyadencsine 5'-

triphosphate (ddATP), 2',3'-dideoxyguanosine 5'-triphosphate (ddGTP), 2',3'-dideoxycytidine 5'-triphosphate (ddCTP), and 2',3'-dideoxythymidine 5'-triphosphate (ddTTP).

In one embodiment, the linker comprises one or more fluorine atoms.

In one embodiment, the linker is selected from the group consisting of:





and

In different embodiments, the linker can comprise a chain structure, or a structure comprising one or more rings, or a structure comprising a chain and one or more rings.

In different embodiments, the linker is cleavable by a means selected from the group consisting of one or more of a physical means, a chemical means, a physical chemical means, heat, and light. In one embodiment, the linker is cleavable by ultraviolet light. In different embodiments, the linker is cleavable by ammonium hydroxide, formamide, or a change in pH (-log H^+ concentration).

In different embodiments of the linker, the chemical moiety comprises biotin, streptavidin or related analogues that have affinity with biotin, phenylboronic acid, salicylhydroxamic acid, an antibody, or an antigen.

In different embodiments, the dideoxynucleotide comprises a cytosine or a thymine with a 5-position, or an adenine or a guanine with a 7-position, and the linker is attached to the 5-position of cytosine or thymine or to the 7-position of adenine or guanine.

The invention provides for the use of any of the linkers described herein in single nucleotide polymorphism detection using mass spectrometry, wherein the linker increases mass separation between different dideoxynucleotides and increases mass spectrometry resolution.

Labeled dideoxynucleotides are provided which comprise a chemical moiety attached via a linker to a 5-position of

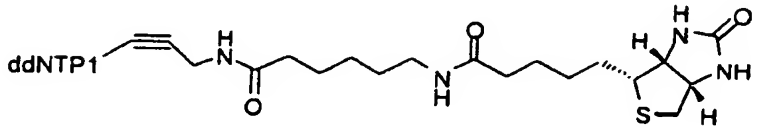
cytosine or thymine or to a 7-position of adenine or guanine.

In one embodiment, the dideoxynucleotide is selected from the group consisting of 2',3'-dideoxyadenosine 5'-triphosphate (ddATP), 2',3'-dideoxyguanosine 5'-triphosphate (ddGTP), 2',3'-dideoxycytidine 5'-triphosphate (ddCTP), and 2',3'-dideoxythymidine 5'-triphosphate (ddTTP).

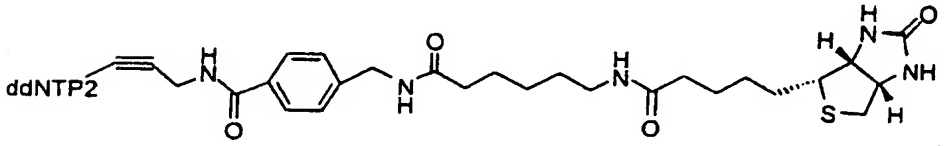
In different embodiments, the linker can comprise a chain structure, or a structure comprising one or more rings, or a structure comprising a chain and one or more rings. In different embodiments, the linker is cleavable by a means selected from the group consisting of one or more of a physical means, a chemical means, a physical chemical means, heat, and light. In one embodiment, the linker is cleavable by ultraviolet light. In different embodiments, the linker is cleavable by ammonium hydroxide, formamide, or a change in pH -log [H^+ concentration].

In different embodiments of the labeled dideoxynucleotide, the chemical moiety comprises biotin, streptavidin, phenylboronic acid, salicylhydroxamic acid, an antibody, or an antigen.

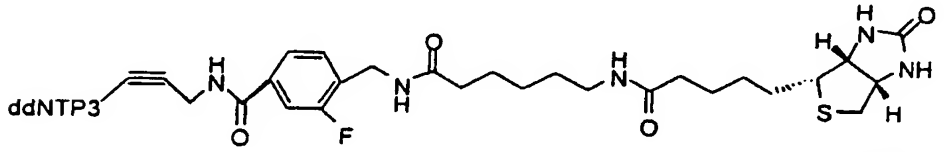
In one embodiment, the labeled dideoxynucleotide is selected from the group consisting of:



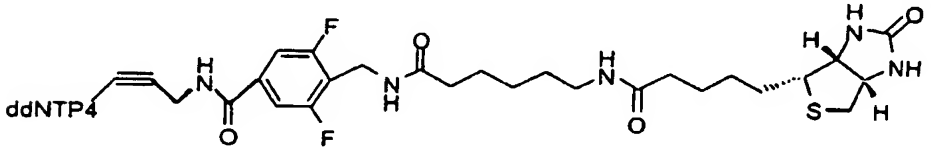
,



,



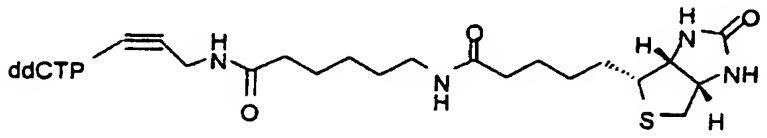
and

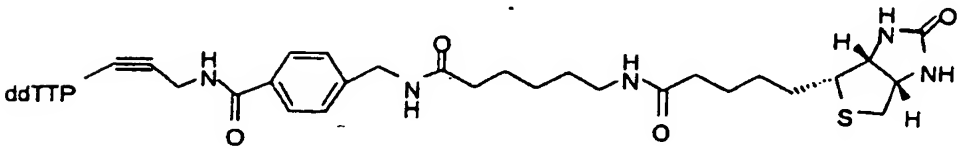


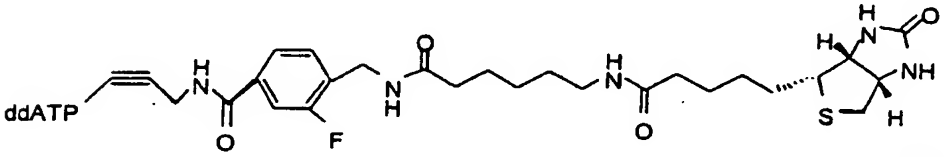
,

wherein ddNTP1, ddNTP2, ddNTP3, and ddNTP4 represent four different dideoxynucleotides, or their analogues.

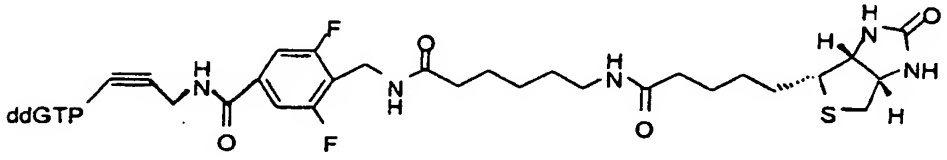
In one embodiment, the labeled dideoxynucleotide is selected from the group consisting of:



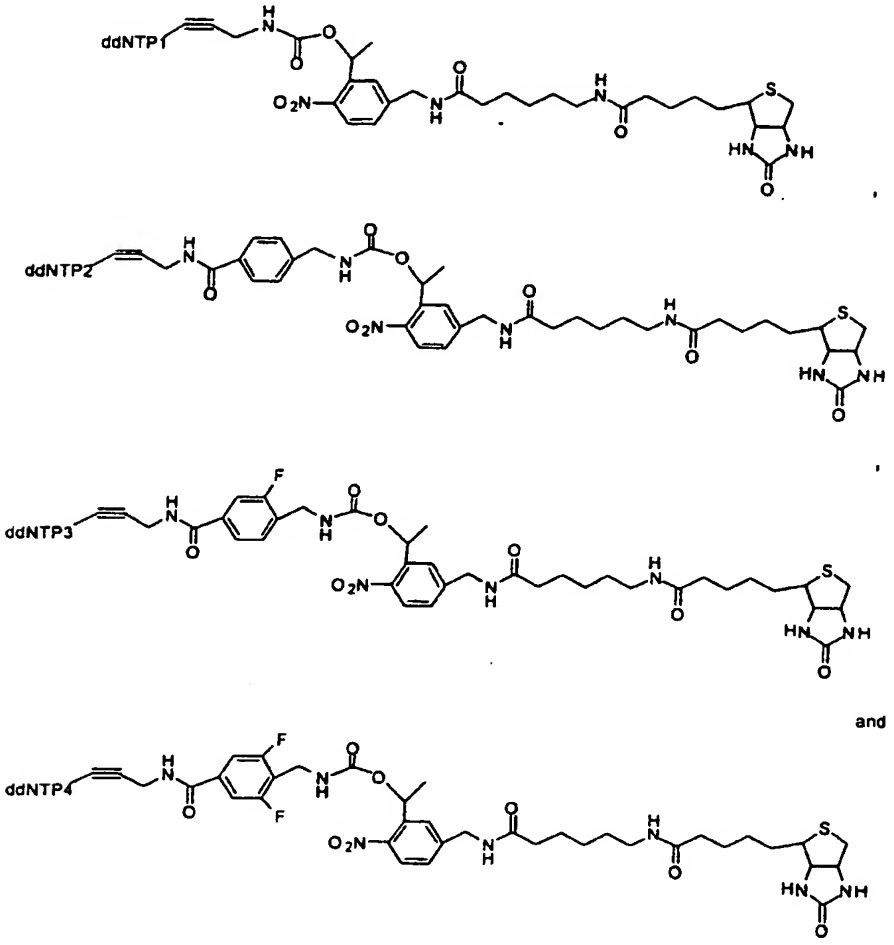




and

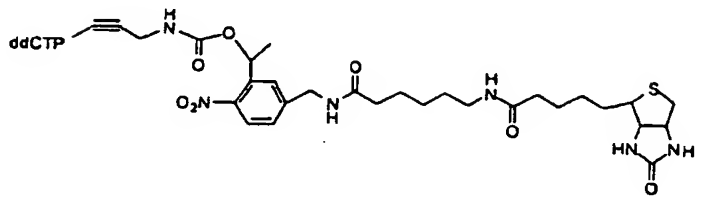


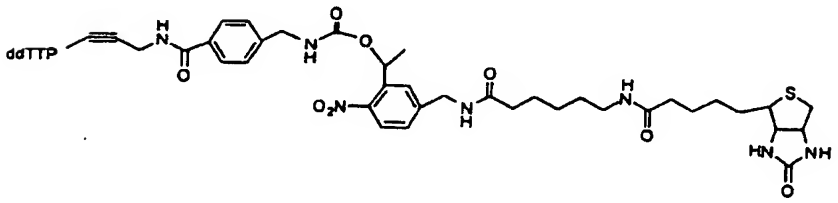
In one embodiment, the labeled dideoxynucleotide is selected from the group consisting of:

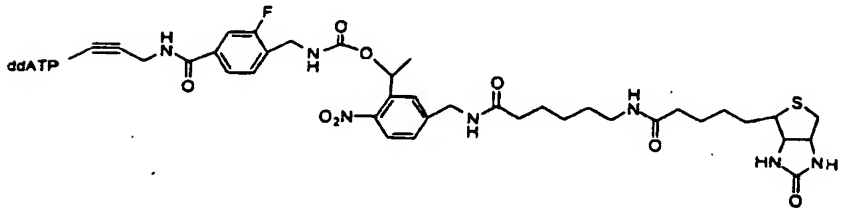


wherein ddNTP1, ddNTP2, ddNTP3, and ddNTP4 represent four different dideoxynucleotides, or their analogues.

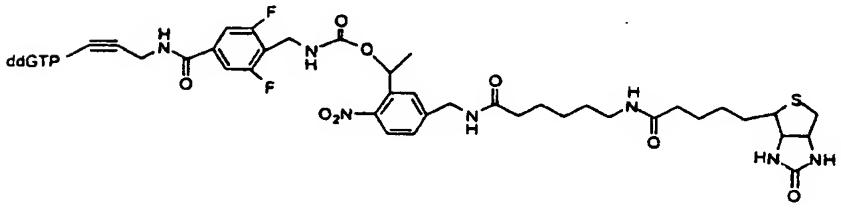
In one embodiment, the labeled dideoxynucleotide is selected from the group consisting of:







and



In one embodiment, the labeled dideoxynucleotide has a molecular weight of 844, 977, 1,017, or 1,051. In one embodiment, the labeled dideoxynucleotide has a molecular weight of 1,049, 1,182, 1,222, or 1,257. Other molecular weights with sufficient mass differences to allow resolution in mass spectrometry can also be used.

In one embodiment the mass spectrometry is matrix-assisted laser desorption/ionization time-of-flight mass spectrometry.

A system is provided for separating a chemical moiety from other components in a sample in solution, which comprises:

(a) a channel coated with a compound that specifically interacts with the chemical moiety at the 3' end of the DNA fragment, wherein the channel comprises a plurality of ends;

(b) a plurality of wells each suitable for holding the sample;

(c) a connection between each end of the channel and a well; and

(d) a means for moving the sample through the channel between wells.

In one embodiment of the system, the interaction between the chemical moiety and the compound coating the surface is a biotin-streptavidin interaction, a phenylboronic acid-salicylhydroxamic acid interaction, or an antigen-antibody interaction.

In one embodiment, the chemical moiety is a biotinylated moiety and the channel is a streptavidin-coated silica

glass channel. In one embodiment, the biotinylated moiety is a biotinylated DNA fragment.

In one embodiment, the chemical moiety can be freed from the surface by disrupting the interaction between the chemical moiety and the compound coating the surface. In different embodiments, the interaction can be disrupted by a means selected from the group consisting of one or more of a physical means, a chemical means, a physical chemical means, heat, and light. In different embodiments, the interaction can be disrupted by ammonium hydroxide, formamide, or a change in pH -log [H^+ concentration].

In one embodiment, the chemical moiety is attached via a linker to another chemical compound. In one embodiment, the other chemical compound is a DNA fragment. In one embodiment, the linker is cleavable by a means selected from the group consisting of one or more of a physical means, a chemical means, a physical chemical means, heat, and light. In one embodiment, the channel is transparent to ultraviolet light and the linker is cleavable by ultraviolet light. Cleaving the linker frees the DNA fragment or other chemical compound from the chemical moiety which remains captured on the surface.

Multi-channel systems are provided which comprise a plurality of any of the single channel systems disclosed herein. In one embodiment, the channels are in a chip. In one embodiment, the multi-channel system comprises 96 channels in a chip. Chips can also be used with fewer or greater than 96 channels.

The invention provides for the use of any of the separation systems described herein for single nucleotide polymorphism detection.

This invention will be better understood from the Experimental Details which follow. However, one skilled in the art will readily appreciate that the specific methods and results discussed are merely illustrative of the invention as described more fully in the claims which follow thereafter.

Experimental Details

Experimental Set I

A. Materials and Methods

PCR amplification. DNA templates containing the polymorphic sites for the human hereditary hemochromatosis gene *HFE* were amplified from genomic DNA in a total volume of 10 µl, that contains 20 ng of genomic DNA, 500 pmol each of forward (C282Y; 5′-CTACCCCCAGAACATCACC-3' (SEQ ID NO: 1), H63D; 5′-GCACTACCTCTTCATGGGTGCC-3' (SEQ ID NO: 2)) and reverse (C282Y; 5′-CATCAGTCACATACCCCA-3' (SEQ ID NO: 3), H63D; 5′-CAGTGAACATGTGATCCCACCC-3' (SEQ ID NO: 4)) primers, 25 µM dNTPs (Amersham Biosciences, Piscataway, NJ), 1 U *Taq* polymerase (Life Technologies, Rockville, MD), and 1× PCR buffer (50 mM KCl, 1.5 mM $MgCl_2$, 10 mM Tris-HCl). PCR amplification reactions were started at 94 °C for 4 min, followed by 45 cycles of 94 °C for 30 s, 59 °C for 30 s and 72 °C for 10 s, and finished with an additicnal extension step of 72 °C for 6 min. Excess primers and dNTPs were degraded by adding 2 U shrimp alkaline phosphatase (Roche Diagnostics, Indianapolis, IN) and *E. Coli* exonuclease I (Boehringer Mannheim, Indianapolis, IN) in 1× shrimp alkaline phosphatase buffer. The reaction mixture was incubated at 37 °C for 45 min followed by enzyme inactivation at 95 °C for 15 min.

Single base extension using biotin-ddNTPs. The synthetic DNA templates containing six nucleotide variations in *p53* gene and the five primers for detecting these variations

are shown in Table 1. These oligonucleotides and an internal mass standard (5'-TTTTTCTTTTTCT-3' (SEQ ID NO: 5), MW = 3855 Da) for MALDI-TOF MS measurement were made using an Expedite nucleic acid synthesizer (Applied Biosystems, Foster City, CA). SBE reactions contained 20 pmol of primer, 10 pmol of biotin-11-ddATP, 20 pmol of biotin-11-ddGTP, 40 pmol of biotin-11-ddCTP (New England Nuclear Life Science, Boston, MA), 80 pmol of biotin-16-ddUTP (Enzo Diagnostics, Inc., Farmingdale, NY), 2 µl Thermo Sequenase reaction buffer, 1 U Thermo Sequenase in its diluted buffer (Amersham Biosciences) and 20 pmol of either synthetic template or 10 µl PCR product in a total reaction volume of 20 µl. For SBE using synthetic template 1, 10 pmol of both wild type and mutated templates were combined with 20 pmol of primers 1 and 3 or 20 pmol of primers 2 and 4. The SBE reaction of primer 5 was performed with template 2 in a separate tube. PCR products from the *HFE* gene were mixed with 20 pmol of the corresponding primers 5'-GGGGAAGAGCAGAGATATACGT-3' (SEQ ID NO: 6) (C282Y) and 5'-GGGGCTCCACACGGCGACTCTC-AT-3' (SEQ ID NO: 7) (H63D) in SBE to detect the two heterozygous genotypes. All extension reactions were thermalcycled for 35 cycles at 94 °C for 10 s and 49 °C for 30 s.

Solid phase purification. 20 µl of the streptavidin-coated magnetic beads (Seradyn, Ramsey, MN) were washed with modified binding and washing (B/W) buffer (0.5 mM Tris-HCl buffer, 2 M NH_4Cl, 1 mM EDTA, pH 7.0) and resuspended in 20 µl modified B/W buffer. Extension reaction mixtures of primers 1-4 with template 1 and primer 5 with template 2 were mixed in a 2:1 ratio, while

extension reaction mixtures from the PCR products of *HFE* gene were mixed in equal amounts. 20 µl of each mixed extension product was added to the suspended beads and incubated for 1 hour. After capture, the beads were washed twice with modified B/W buffer, twice with 0.2 M triethyl ammonium acetate (TEAA) buffer and twice with deionized water. The primer extension products were released from the magnetic beads by treatment with 8 µl 98 % formamide solution containing 2 % 0.2 M TEAA buffer at 94 °C for 5 min. The released primer extension products were precipitated with 100 % ethanol at 4 °C for 30 min, and centrifuged at 4 °C and 14000 RPM for 35 min.

MALDI-TOF MS analysis. The purified primer extension products were dried and re-suspended in 1 µl deionized water and 2 µl matrix solution. The matrix solution was made by dissolving 35 mg of 3-hydroxypicolinic acid (3-HPA; Aldrich, Milwaukee, WI) and 6 mg of ammonium citrate (Aldrich) in 0.8 ml of 50 % acetonitrile. 10 pmol internal mass standard in 1 µl of 50 % acetonitrile was then added to the sample. 0.5 µl of this mixture containing the primer extension products and internal standard was spotted on a stainless steel sample plate, air-dried and analyzed using an Applied Biosystems Voyager DE Pro MALDI-TOF mass spectrometer. All measurements were taken in linear positive ion mode with a 25 kV accelerating voltage, a 94 % grid voltage and a 350 ns delay time. The obtained spectra were processed using the Voyager data analysis package.

B. Detection of Single Nucleotide Polymorphism Using Biotinylated Dideoxynucleotides and Mass Spectrometry

Solid phase capturable biotinylated dideoxynuclectides (biotin-ddNTPs) were used in single base extension for multiplex genotyping by mass spectrometry (MS). In this method, oligonucleotide primers that have different molecular weights and that are specific to the polymorphic sites in the DNA template are extended with biotin-ddNTPs by DNA polymerase to generate 3'-biotinylated DNA extension products (Figure 1). These products are then captured by streptavidin-coated solid phase magnetic beads, while the unextended primers and other components in the reaction are washed away. The pure extension DNA products are subsequently released from the solid phase and analyzed with matrix-assisted laser desorption/ionization time-of-flight MS. The mass of the extension DNA products is determined using a stable oligonucleotide as a common internal mass standard. Since only the pure extension DNA products are introduced to MS for analysis, the resulting mass spectrum is free of non-extended primer peaks and their associated dimers, which increases the accuracy and scope of multiplexing in single nucleotide polymorphism (SNP) analysis. The solid phase purification approach also facilitates desalting of the captured oligonucleotides, which is essential for accurate mass measurement by MS.

Four biotin-ddNTPs with distinct molecular weights were selected to generate extension products that have a two-fold increase in mass difference compared to that with conventional ddNTPs. This increase in mass difference

provides improved resolution and accuracy in detecting heterozygotes in the mass spectrum.

The "lock and key" functionality of biotin and streptavidin is often utilized in biological sample preparation as a way to remove undesired impurities (23). In different embodiments of the methods described herein, affinity systems other than biotin-streptavidin can be used. Such affinity systems include but are not limited to phenylboronic acid-salicylhydroxamic acid (31) and antigen-antibody systems.

The multiplex genotyping approach was validated by detecting six nucleotide variations from synthetic DNA templates that mimic mutations in exons 7 and 8 of the p53 gene. Sequences of the templates and the corresponding primers are shown in Table 1 along with the masses of the primers and their extension products. The mass increase of the resulting single base extension products in comparison with the primers is 665 Da for addition of biotin-ddCTP, 688 Da for addition of biotin-ddATP, 704 Da for addition of biotin-ddGTP and 754 Da for addition of biotin-ddUTP. The mass data in Table 1 indicate that the smallest mass difference among any possible extensions of a primer is 16 Da (between biotin-ddATP and biotin-ddGTP). This is a substantial increase over the smallest mass difference between extension products using standard ddNTPs (9 Da between ddATP and ddTTP). This mass increase yields improved resolution of the peaks in the mass spectrum. Increased mass difference in ddNTPs fosters accurate detection of heterozygous genotypes (15), since an A/T heterozygote with a mass difference of 9 Da using conventional ddNTPs

can not be well resolved in the MALDI-TOF mass spectra. The five primers for each polymorphic site were designed to produce extension products without overlapping masses. Primers extended by biotin-ddNTPs were purified and analyzed by MALDI-TOF MS according to the scheme in Figure 1. Extension products of all five primers were well-resolved in the mass spectrum free from any unextended primers (Figure 2A), allowing each nucleotide variation to be unambiguously identified. Unextended primers occupy the mass range in the mass spectrum decreasing the scope of multiplexing, and excess primers can dimerize to form false peaks in the mass spectrum (21). The excess primers and their associated dimers also compete for the ion current, reducing the detection sensitivity of MS for the desired DNA fragments. These complications were completely removed by carrying out SBE using biotin-ddNTPs and solid phase capture. Extension products for all four biotin-ddNTPs were clearly detected with well resolved mass values. The relative masses of the primer extension products in comparison to the internal mass standard revealed the identity of each nucleotide at the polymorphic site. In the case of heterozygous genotypes, two peaks, one corresponding to each allele (C/A), are clearly distinguishable in the mass spectrum shown in Figure 2A.

Table 1. Oligonucleotide primers and synthetic DNA templates for detecting mutations in the p53 gene. (Top) The sequences and the calculated masses of primers and the four possible single base extension products relative to the internal mass standard are listed. The bold numbers refer to the nucleotide variations detected in the p53 gene. (Bottom) The six nucleotide variations in template 1 and 2 are shown in bold letters. Template 1 contains a heterozygous genotype (G/T). Primers 1-5 = SEQ ID NOs: 8-12, respectively.

Primers	Primer sequences	Masses (Da)	Masses of single base extension products (Da)			
			Biotin-ddCTP *Δ665*	Biotin-ddATP *Δ688*	Biotin-ddGTP *Δ704*	Biotin-ddUTP *Δ754*
1	5'-AGAGGATCCAACCGAGAC-3'	1656	2321	2344	2360	**2410**
2	5'-TGGTGGTAGGTGATGTTGATGTA-3'	3350	**4015**	4038	4054	4103
3	5'-CACATTGTCAAGGACGTACCCG-3'	2833	**3498**	3521	3538	3587
4	5'-TACCCGCCGTACTTGGCCTC-3'	2134	**2799**	**2822**	2838	2480
5	5'-TCCACGCACAAACACGGACAG-3'	2507	3172	3195	**3211**	3261

Templates	Template sequences
1	5'-TACCCG/TGAGGCCAAGTACGGCGGGTACGTCCTTGACAATGTGTACATCAACATCACCTACCACCATGT
2	CAGTCTCGGTTGGATCCTCTATTGTGTCCGGG-3' (SEQ ID NO: 13)
	5'-GAAGGAGACACGCGGCCAGAGAGGGTCCTGTCCGTGTTTGTGCGTGGAGTTTCGACAAGGCAGGGTCAT
	CTAATGGTGATGAGTCCTATCCTTTTCTCTTCGTTCTCCGT-3' (SEQ ID NO: 14)

One advantage of MALDI-TOF MS in comparison to other detection techniques is its ability to simultaneously measure masses of DNA fragments over a certain range.

In order to explore this feature to detect multiple SNPs in a single spectrum, if unextended primers are not removed, masses of all primers and their extension products must have sufficient differences to yield adequately resolved peaks in the mass spectrum. Ross et al. simultaneously detected multiple SNPs by carefully tuning the masses of all primers and extension products so that they would lie in the range of 4.5 kDa and 7.6 kDa without overlapping (14). Since the unextended primers occupy the mass range in the mass spectrum, by eliminating them, the approach disclosed herein will significantly increase the scope of multiplexing in SNP analysis.

To demonstrate the ability of this method to discriminate SNPs in genomic DNA, two disease associated SNPs were genotyped in the human hereditary hemochromatosis (HHC) gene HFE. HHC is a common genetic condition in Caucasians with approximately 1/400 Caucasians homozygous for the C282Y mutation leading to iron overload and potentially liver failure, diabetes and depression (22). A subset of individuals who are compound heterozygotes for the C282Y and H63D mutations also manifest iron overload. Because of the high prevalence of these mutations and the ability to prevent disease manifestations by phlebotomy, accurate methods for genotyping these two SNPs will foster genetic screening for this condition. Two PCR products were generated from human genomic DNA for the C282Y and H63D polymorphic

sites of the HFE gene and then used these products for SBE with biotin-ddNTPs. After the extension reaction, products were purified using solid phase capture according to the scheme in Figure 1 and analyzed by MALDI-TOF MS. The mass spectrum obtained from this experiment is shown in Figure 2B. Extension products of each primer were readily identified by their mass relative to the internal mass standard. Heterozygous genotypes of A/G and C/G with a mass difference of 16 Da and 39 Da respectively were accurately detected at the C282Y and H63D polymorphic sites.

These results indicate that the use of solid phase capturable biotin-ddNTPs in SBE, coupled with MALDI-TOF MS detection, provides a rapid and accurate method for multiplex SNP detection over broad mass ranges and should greatly increase the number of SNPs that can be detected simultaneously. In multiplex SBE reactions, the oligonucleotide primers and their dideoxynucleotide extension products differ by only one base pair, which requires analytical techniques with high resolution to resolve. In addition, a primer designed to detect one polymorphism and an extension product from another polymorphic site may have the same size, which can not be separated by electrophoresis and other conventional chromatographic or size exclusion methods. Methods for purifying DNA samples using the strong interaction of biotin and streptavidin are widely used (23-27). By introducing the biotin moiety at the 3′ end of DNA, the solid phase based affinity purification approach described here is a unique and effective method to remove the oligonucleotide primers from the dideoxynucleotide extension products.

To increase the stability of DNA fragments for MALDI-TOF MS measurement in multiplex SNP analysis, nucleotide analogues (28) and peptide nucleic acid (9) can be used in the construction of the oligonucleotide primers. It has been shown that MALDI-TOF MS could detect DNA fragments up to 100 bp with sufficient resolution (29). The mass difference between each adjacent DNA fragment is approximately 300 Da. Thus, with a mass difference of 100 Da for each primer in designing a multiplex SNP analysis project, at least 300 SNPs can be analyzed in a single spot of the sample plate by MS. It is a routine method now to place 384 spots in each sample plate in MS analysis. Thus, each plate can produce over 100,000 SNPs, which is roughly the entire SNPs in all the coding regions of the human genome. This level of multiplexing should be achievable by mass tagging the primers with stable chemical groups in SBE using biotin-ddNTPs. For SNP sites of interest, a master database of primers and the resulting masses of all four possible extension products can be constructed. The experimental data from MALDI-TOF MS can then be compared with this database to precisely identify the library of SNPs automatically. This method coupled with future improvements in mass spectrometer detector sensitivity (30) will provide a platform for high-throughput SNP identification unrivaled in speed and accuracy.

C. Design and Synthesis of Biotinylated dideoxynucleotides with Mass Tags

The ability to distinguish various bases in DNA using mass spectrometry is dependent on the mass differences of the bases in the spectra. For the above work, the

smallest difference in mass between any two nucleotides is 16 daltons (see Table 1). Fei et al. (15) have shown that using dye-labeled ddNTP paired with a regular dNTP to space out the mass difference, an increase in the detection resolution in a single nucleotide extension assay can be achieved. To enhance the ability to distinguish peaks in the spectra, the current application discloses systematic modification of the biotinylated dideoxynucleotides by incorporating mass linkers assembled using 4-aminomethyl benzoic acid derivatives to increase the mass separation of the individual bases. The mass linkers can be modified by incorporating one or two fluorine atoms to further space out the mass differences between the nucleotides. The structures of four biotinylated ddNTPs are shown in Figure 3. ddCTP-11-biotin is commercially available (New England Nuclear, Boston). ddTTP-Linker I-11-Biotin, ddATP-Linker II-11-Biotin and ddGTP-Linker III-11-Biotin are synthesized as shown, for example, for ddATP-Linker II-11-Biotin in Figure 5. In designing these mass tag linker modified biotinylated ddNTPs, the linkers are attached to the 5-position on the pyrimidine bases (C and T), and to the 7-position on the purines (A and G) for subsequent conjugation with biotin. It has been established that modification of these positions on the bases in the nucleotides, even with bulky energy transfer fluorescent dyes, still allows efficient incorporation of the modified nucleotides into the DNA strand by DNA polymerase (32, 33). Thus, the ddNTPs-Linker-11-biotin can be incorporated into the growing strand by the polymerase in DNA sequencing reactions.

Larger mass separations will greatly aid in longer read lengths where signal intensity is smaller and resolution is lower. The smallest mass difference between two individual bases is over three times as great in the mass tagged biotinylated ddNTPs compared to normal ddNTPs and more than double that achieved by the standard biotinylated ddNTPs as shown in Table 2.

Table 2. Relative mass differences (daltons) of dideoxynucleotides using ddCTP as a reference.

Base	Standard ddNTP	Commercial Biotinylated ddNTP	Biotinylated ddNTP with mass tag linker
C relative to C	0	0	0 (no linker)
T relative to C	15	89 (16 linker)	125 (Linker I)
A relative to C	24	24	165 (Linker II)
G relative to C	40	40	200 (Linker III)
Smallest relative difference	9	16	35

Three 4-aminomethyl benzoic acid derivatives **Linker I**, **Linker II** and **Linker III** are designed as mass tags as well as linkers for bridging biotin to the corresponding dideoxynucleotides. The synthesis of **Linker II** (Figure 4) is described here to illustrate the synthetic procedure. 3-Fluoro-4-aminomethyl benzoic acid that can be easily prepared via published procedures (41, 42) is first protected with trifluoroacetic anhydride, then converted to N-hydroxysuccinimide (NHS) ester with disuccinimidylcarbonate in the presence of diisopropylethylamine. The resulting NHS ester is subsequently coupled with commercially available

propargylamine to form the desired compound, **Linker II**. Using an analogous procedure, **Linker I** and **Linker III** can be easily constructed.

Figure 5 describes the scheme required to prepare biotinylated ddATP-Linker II-11-Biotin using well-established procedures (34-36). 7-I-ddA is coupled with linker II in the presence of tetrakis(triphenylphosphine) palladium(0) to produce 7-Linker II-ddA, which is phosphorylated with $POCl_3$ in butylammonium pyrophosphate (37). After removing the trifluoroacetyl group with ammonium hydroxide, 7-Linker II-ddATP is produced, which then couples with sulfo-NHS-LC-Biotin (Pierce, Rockford IL) to yield the desired ddATP-Linker II-11-Biotin. Similarly, ddTTP-Linker I-11-Biotin, and ddGTP-Linker III-11-Biotin can be synthesized.

D. Design and Synthesis of Mass Tagged ddNTPs Containing Photocleavable Biotin

A schematic of capture and cleavage of the photocleavable linker on the streptavidin coated porous surface is shown in Figure 6. At the end of the reaction, the reaction mixture consists of excess primers, enzymes, salts, false stops, and the desired DNA fragment. This reaction mixture is passed over a streptavidin-coated surface and allowed to incubate. The biotinylated fragments are captured by the streptavidin surface, while everything else in the mixture is washed away. Then the fragments are released into solution by cleaving the photocleavable linker with near ultraviolet (UV) light, while the biotin remains attached to the streptavidin that is covalently bound to the surface. The pure DNA fragments can then be

crystallized in matrix solution and analyzed by mass spectrometry. It is advantageous to cleave the biotin moiety since it contains sulfur which has several relatively abundant isotopes. The rest of the DNA fragments and linkers contain only carbon, nitrogen, hydrogen, oxygen, fluorine and phosphorous, whose dominant isotopes are found with a relative abundance of 99% to 100%. This allows high resolution mass spectra to be obtained. The photocleavage mechanism (38, 39) is shown in Figure 7. Upon irradiation with ultraviolet light at 300-350 nm, the light sensitive o-nitroaromatic carbonamide functionality on DNA fragment 1 is cleaved, producing DNA fragment 2, PC-biotin and carbon dioxide. The partial chemical linker remaining on DNA fragment 2 is stable for detection by mass spectrometry.

Four new biotinylated ddNTPs disclosed here, ddCTP-PC-Biotin, ddTTP-Linker I-PC-Biotin, ddATP-Linker II-PC-Biotin and ddGTP-Linker III-PC-Biotin are shown in Figure 8. These compounds are synthesized by a similar chemistry as shown for the synthesis of ddATP-Linker II-11-Biotin in Figure 6. The only difference is that in the final coupling step NHS-PC-LC-Biotin (Pierce, Rockford IL) is used, as shown in Figure 9. The photocleavable linkers disclosed here allow the use of solid phase capturable terminators and mass spectrometry to be turned into a high throughput technique for DNA analysis.

E. Overview of capturing a DNA fragment terminated with a ddNTP on a surface and freeing the ddNTP and DNA fragment

The DNA fragment is terminated with a dideoxynucleoside monophosphate (ddNMP). The ddNMP is attached via a linker to a chemical moiety ("X" in Figure 10). The DNA fragment terminated with ddNMP is captured on the surface through interaction between chemical moiety "X" and a compound on or attached to the surface ("Y" in Figure 10). The present application discloses two methods for freeing the captured DNA fragment terminated with ddNMP. In the situation illustrated in the lower part of Figure 10, the DNA fragment terminated with ddNMP is freed from the surface by disrupting or breaking the interaction between chemical moiety "X" and compound "Y". In the upper part of Figure 10, the DNA fragment terminated with ddNMP is attached to chemical moiety "X" via a cleavable linker which can be cleaved to free the DNA fragment terminated with ddNMP.

Different moieties and compounds can be used for the "X" - "Y" affinity system, which include but are not limited to, biotin-streptavidin, phenylboronic acid-salicylhydroxamic acid (31), and antigen-antibody systems.

In different embodiments, the cleavable linker can be cleaved and the "X" - "Y" interaction can be disrupted by a means selected from the group consisting of one or more of a physical means, a chemical means, a physical chemical means, heat, and light. In one embodiment, ultraviolet light can be used to cleave the cleavable linker. Chemical means include, but are not limited to,

ammonium hydroxide (40), formamide, or a change in pH (-log H^+ concentration) of the solution.

F. High density streptavidin-coated, porous silica channel system.

Streptavidin coated magnetic beads are not ideal for using the photocleavable biotin capture and release process for DNA fragments, since they are not transparent to UV light. Therefore, the photocleavage reaction is not efficient. For efficient capture of the biotinylated fragments, a high-density surface coated with streptavidin is essential. It is known that the commercially available 96-well streptavidin coated plates cannot provide a sufficient surface area for efficient capture of the biotinylated DNA fragments. Disclosed in this application is a porous silica channel system designed to overcome this limitation.

To increase the surface area available for solid phase capture, porous channels are coated with a high density of streptavidin. For example, ninety-six (96) porous silica glass channels can be etched into a silica chip (Figure 11). The surfaces of the channels are modified to contain streptavidin as shown in Figure 12. The channel is first treated with 0.5 M NaOH, washed with water, and then briefly pre-etched with dilute hydrogen fluoride. Upon cleaning with water, the capillary channel is coated with high density 3-aminopropyltrimethoxysilane in aqueous ethanol (43). An excess of disuccinimidyl glutarate in N,N-dimethylformamide (DMF) is then introduced into the capillary to ensure a highly efficient conversion of the surface end group to a

succinimidyl ester. Streptavidin is then conjugated with the succinimidyl ester to form a high-density surface using excess streptavidin solution. The resulting 96-channel chip is used as a purification cassette.

A 96-well plate that can be used with biotinylated terminators for DNA analysis is shown in Figure 11. In the example shown, each end of a channel is connected to a single well. However, for other applications, the end of a channel could be connected to a plurality of wells. Pressure is applied to drive the samples through a glass capillary into the channels on the chip. Inside the channels the biotin is captured by the covalently bound streptavidin. After passing through the channel, the sample enters into a clean plate in the other end of the chip. Pressure applied in reverse drives the sample through the channel multiple times and ensures a highly efficient solid phase capture. Water is similarly added to drive out the reaction mixture and thoroughly wash the captured fragments. After washing, the chip is irradiated with ultraviolet light to cleave the photosensitive linker and release the DNA fragments. The fragment solution is then driven out of the channel and into a collection plate. After matrix solution is added, the samples are spotted on a chip and allowed to crystallize for detection by MALDI-TOF mass spectrometry. The purification cassette is cleaned by chemically cleaving the biotin-streptavidin linkage, and is then washed and reused.

Experimental Set II

A. Synopsis

The following experiments show the simultaneous genotyping of 30 nucleotide variations in the *p53* gene from human tumors in one tube, by using solid phase capturable dideoxynucleotides to generate single base extension products which are detected by mass spectrometry. Both homozygous and heterozygous genotypes are accurately determined with digital resolution. This is the highest level of SNP multiplexing reported thus far using mass spectrometry, indicating the approach will have wide applications in screening a repertoire of genotypes in candidate genes as potential markers for cancer and other diseases.

B. Introduction

With the completion of the Human Genome Project, a stage has been set to screen genetic mutations for identifying disease genes in a genomewide scale (44). Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), which allows rapid DNA sample measurement yielding digital data, has been explored to detect single nucleotide polymorphisms (SNPs) using invasive cleavage (11) and primer-directed base extension (14, 45). Conventional single base extension (SBE) methods using MS to measure multiplex SNPs require unambiguous simultaneous detection of a library of primers and their extension products. However, limitations in resolution and sensitivity of MALDI-TOF MS

for longer DNA molecules make it difficult to simultaneously measure DNA fragments over a large mass range. The requirement to measure both primers and their extension products in this range limits the scope of multiplexing. The use of MALDI-TOF MS and molecular affinity for multiplex digital SNP detection using solid phase capturable (SPC) dideoxynucleotides and SBE has recently been explored, establishing the feasibility of simultaneously measuring 20 SNPs in synthetic DNA templates (46). This study shows the simultaneous genotyping of 30 nucleotide variations, corresponding to known sites of cancer-associated somatic mutations, in exons 5, 7 and 8 of the *p53* gene from human tumors in one tube using the SPC-SBE method. This is the highest level of multiplexing reported thus far using mass spectrometry for SNP analysis.

C. Materials and Methods

Multiplex PCR and single base extension reactions Multiplex PCR was performed to amplify 3 regions in exons 5, 7 and 8 of the *p53* gene. The primers for each region were 5'-TATCTGTTCACTTGTGCCC-3' (exon 5, forward), 5'-CAGAGGCCTGGGGA-CCCTG-3' (exon 5, reverse), 5'-CTGCTTGCCACAGGTCTC-3' (exon 7, forward), 5'-CACAGCAG-GCCAGTGTGC-3' (exon 7, reverse), 5'-GGACCTGATTTCCTTAC-TG-3' (exon 8, forward), and 5'-TGAATCTGAGGCATAACTG-3' (exon 8, reverse). The 45 l PCR reaction consisted of 180 ng genomic DNA, 1.5 nmol dNTP, 4.5 l 10X PCR buffer, 15 mM $MgCl_2$, 4 pmol of forward and reverse primers for exons 5 and 7, 6 pmol of forward and reverse primers for exon 8, and 1.0 U of JumpStart RedAccuTaq DNA Polymerase. After

a 5 min 96 °C hot start, the touchdown PCR program was performed with 10 cycles of 96 °C (30 sec), 67 °C to 57 °C (-1.0 °C per cycle, 30 sec) and 72 °C (30 sec), an additional 30 cycles of 96 °C (30 sec), 57 °C (30 sec) and 72 °C (30 sec), and a final extension at 72 °C for 7 min. The 30 SBE primers (Table 3) were designed to yield extension products with a sufficient mass difference and to be extended simultaneously in a single tube. Primer sequences were designed to avoid any overlap in mass, and the formation of secondary structures. To evenly separate the masses of such a large number of primers for SBE, some primers were synthesized using methyl-dC and dU phosphoramidites (Glen Research) to replace dC and dT respectively. Substitution of dC by methyl-dC increased the primer mass by 14 Da whereas a change from dT to dU decreased the mass by 14 Da. Primers were synthesized using an Applied Biosystems DNA synthesizer. The procedures for the SBE, solid phase purification and MALDI-TOF MS measurement were performed as described (Kim *et al.*, *Analytical Biochemistry* 2003, 316, 251). Direct DNA sequencing was conducted using energy transfer terminator chemistry and a MegaBACE 1000 capillary DNA sequencer (Amersham Bioscience).

D. Discussion

Thirty polymorphic sites, including the most frequently mutated *p53* codons, were chosen to explore the high multiplexing scope of the SPC-SBE method (Figure 1). Thirty primers specific to each polymorphic site were designed to yield SBE products with sufficient mass differences. This was achieved by tuning the mass of some primers using methyl-dC and dU to replace dC and dT,

respectively. Human genomic DNA was amplified by multiplex PCR to produce amplicons of three *p53* exons. The 30 primers were mixed with the PCR products and biotinylated dideoxynucleotides for SBE to generate 3'-biotinylated extension DNA products. These products were then captured by streptavidin-coated solid phase magnetic beads, while the unextended primers and other components in the reaction were washed away. The pure DNA products were subsequently released from the solid phase and analyzed by MALDI-TOF MS. The nucleotide at the polymorphic site is accurately identified by the mass of the DNA extension product in a mass spectrum. Since only the DNA extension products are isolated for MS analysis, the resulting mass spectrum is free of non-extended primer peaks and their associated dimers, increasing accuracy and scope of multiplexing. The solid phase purification also facilitates desalting of the captured DNA, a process that is critical for accurate mass measurement by MALDI-TOF MS.

The SPC-SBE genotyping approach was used to analyze nucleotide variations in 30 codons of 3 exons of the *p53* gene from 30 Wilms' tumors, 19 head and neck squamous carcinomas and 3 colorectal carcinomas. Primer sequences are shown in Table 3 along with the masses of the primers and their extension products. Extension products of all 30 primers were resolved in the mass spectrum, free from any unextended primers, yielding digital data to unambiguously determine each nucleotide variation (Figures 13A-13C). Unextended primers occupy the mass range in the mass spectrum decreasing the scope of multiplexing, and excess primers can dimerize to form false peaks in the mass spectrum (21). The excess primers

and their associated dimers also compete for the ion current, reducing the detection sensitivity of MS for the desired DNA fragments. These complications were completely removed in the SPC-SBE method. When using conventional ddNTPs, the mass difference between ddATP and ddTTP is 9 Da, which is difficult to resolve by MALDI-TOF MS (15). In the SPC-SBE method using biotinylated ddNTPs, the difference between A and T is increased to 66 Da, which fosters accurate detection of heterozygous genotypes.

None of the 30 Wilms' tumor samples showed somatic mutations for the 30 polymorphic sites tested, yielding 30 distinct peaks corresponding to the wild type *p53* sequences in a mass spectrum (Figure 13A). In contrast, two of the 19 head and neck tumor samples contained a genetic variation; one at codon 157 (G/T heterozygous configuration; primary tumor biopsy; Figure 13B) and the other at codon 151 (C to T homozygous; squamous carcinoma cell line; Figure 14). In the three colorectal tumor cell lines tested, one (HCT-116) had 30 wild type *p53* sequences for the 30 sites, yielding a mass spectrum similar to the one shown in Figure 13A, while the other two (HT-29 and SW-480) had a G to A homozygous mutation in codon 273 (Figure 13C). Both heterozygous and homozygous genotypes were clearly detected in the 30 codons with great accuracy. The G/T heterozygote (4684/4734 Da) was shown with two peaks corresponding to the wild type and mutant alleles, respectively (Figure 13B). These data, confirmed by direct DNA sequencing, are consistent with the known paucity of the *p53* mutations in Wilms' tumor, and the known occurrence of such mutations in squamous carcinomas and colorectal

carcinomas.

It has been reported that MALDI-TOF MS could detect DNA sequencing fragments up to 100 bp with sufficient resolution using cleavable primers (29). The mass difference between each adjacent DNA sequencing fragment is approximately 300 Da. In principle, with a mass difference of 100 Da for each primer in designing a multiplex SNP analysis project using the SPC-SBE method, at least 300 SNPs can be analyzed in a single spot of an MS sample plate. Thus, each MS sample plate with 384 spots can produce over 100,000 SNPs, which is roughly the number of tag SNPs required to identify all the haplotypes in the human genome. This level of multiplexing should be achievable by mass tuning the primers with nucleotide analogues containing stable chemical groups (28). It is anticipated that the SPC-SBE high-throughput digital SNP detection approach will have wide applications in screening a repertoire of genotypes in candidate genes as potential markers for cancer and other diseases.

Table 3. Thirty *p53* codons and the corresponding 30 SBE primers. The position of the nucleotide variation tested in each codon is shown in bold. The primer sequence and modification is specified and the modified nucleotides are shown in bold. The mass of each primer is indicated along with the mass of all four possible SBE products. The mass values in bold specify the wild type nucleotide sequences (ddNTP-B = Biotinylated dideoxynucleotides).

Primer Number	Exon	Codon	Sequences (5'-3')	Modification	Primer	Mass of Single Base Extention Products (D			
					Mass (Da)	ddATP-B	ddCTP-B	ddGTP-B	ddUTP-B
1	5	179 (CAT)	GCGCTGCCCCCAC	None	3857	4545	4522	4561	4611
2	5	157 (GTC)	GCCCGGCACCCGC	methyl C	3980	4668	4645	4684	4734
3	5	179 (CAT)	GCGCTGCCCCCACC	None	4146	4834	4811	4850	4900
4	5	163 (TAC)	CGCCATGGCCATCT	methyl C	4270	4958	4935	4974	5024
5	5	158 (CGC)	CCGGCACCCGCGTCC	None	4475	5163	5140	5179	5229
6	7	248 (CGG)	TGGGCGGCATGAACC	None	4618	5306	5283	5322	5372
7	5	132 (AAG)	TCCCCTGCCCTCAACA	methyl C	4736	5424	5401	5440	5490
8	8	298 (GAG)	AGGGGAGCCTCACCAC	None	4876	5564	5541	5580	5630
9	8	285 (GAG)	GAGAGACCGGCGCACA	methyl C	4995	5683	5660	5699	5749
10	5	161 (GCC)	CCCGCGTCCGCGCCATG	None	5108	5796	5773	5812	5862
11	7	249 (AGG)	GGCGGCATGAACCGGAG	methyl C	5341	6029	6006	6045	6095
12	8	266 (GGA)	GTAGTGGTAATCTACTGG	dU	5486	6174	6151	6190	6240
13	8	286 (GAA)	AGAGACCGGCGCACAGAG	methyl C	5638	6326	6303	6342	6392
14	7	258 (GAA)	CCTCACCATCATCACACTG	methyl C	5765	6453	6430	6469	6519
15	5	176 (TGC)	ACGGAGGTTGTGAGGCGCT	dU	5897	6585	6562	6601	6651
16	5	152 (CCG)	GTGGGTTGATTCCACACCCC	dU	6041	6729	6706	6745	6795
17	8	273 (CGT)	ACGGAACAGCTTTGAGGTGC	None	6182	6870	6847	6886	6938
18	7	234 (TAC)	CTGACTGTACCACCATCCACT	None	6286	6974	6951	6990	7040
19	7	248 (CGG)	TCCTGCATGGGCGGCATGAAC	dU	6405	7093	7070	7109	7159
20	7	249 (AGG)	GCATGGGCGGCATGAACCGGA	None	6521	7209	7186	7225	7275
21	8	282 (CGG)	TTGTGCCTGTCCTGGGAGAGAC	dU	6698	7386	7363	7402	7452
22	8	278 (CCT)	TGAGGTGCGTGTTTGTGCCTGT	None	6819	7507	7484	7523	7573
23	5	135 (TGC)	CCCTGCCCTCAACAAGATGTTTT	None	6935	7623	7600	7639	7689
24	7	245 (GGC)	TGTGTAACAGTTCCTGCATGGGC	dU	7043	7731	7708	7747	7797
25	7	237 (ATG)	TACCACCATCCACTACAACTACAT	None	7170	7858	7835	7874	7924
26	7	242 (TGC)	ACAACTACATGTGTAACAGTTCCT	dU	7282	7970	7947	7986	8036
27	7	241 (TCC)	ACTACAACTACATGTGTAACAGTT	methyl C	7390	8078	8055	8094	8144
28	8	275 (TGT)	GGAACAGCTTTGAGGTGCGTGTTT	methyl C	7497	8185	8162	8201	8251
29	5	141 (TGC)	ATGTTTTGCCAACTGGCCAAGACCT	None	7617	8305	8282	8321	8371
30	5	175 (CGC)	CAGCACATGACGGAGGTTGTGAGGC	None	7772	8460	8437	8476	8526

REFERENCES

1) Kwok, P.-Y. (2000) Pharmacogenomics 1, 95-100.

2) Roses A. (2000) Pharmacogenetics and the practice of medicine. *Nature*. 405: 857-865.

3) The International SNP Map Working Group (2001) Nature 409, 928-933.

4) Beavis, R. C. & Chait, B. T. (1989) Rapid Commun. Mass Spectrom. 3, 436-439.

5) Li, J., Butler, J. M., Tan, Y., Lin, H., Royer, S., Ohler, L., Shaler, T. A., Hunter, J. A., Pollart, D. J., Monforte, J. A. & Becker, C. H. (1999) Electrophoresis 20, 1258-1265.

6) Griffin, T. J. & Smith, L. M. (2000) Trends. Biotechnol. 18, 77-84.

7) Graber, J. H., Smith, C. L. & Cantor, C. R. (1999) Genetic Analysis: Biomol. Eng. 14, 215-219.

8) Stoerker, J., Mayo, J. D., Tetzlaff, C. N., Sarracino, D. A., Schwope, I. & Richert, C. (2000) Nat. Biotechnol. 18, 1213-1216.

9) Ross, P. L., Lee, K. & Belgrader, P. (1997) Anal. Chem. 69, 4197-4202.

10) Jiang-Baucom, P., Girard, J. E., Butler, J. & Belgrader, P. (1997) Anal. Chem. 69, 4894-4898.

11) Griffin, T. J., Hall, J. G., Prudent, J. R. & Smith, L. M. (1999) Proc. Natl. Acad. Sci. USA. 96, 6301-6306.

12) Lyamichev, V., Mast, A. L., Hall, J. G., Prudent, J. R., Kaiser, M. W., Takova, T., Kwiatkowski, R. W., Sander, T. J., de Arruda, M., Arco, D. A., Neri, B. P. & Brow, M. A. D. (1999) Nat. Biotechnol. 17, 292-296.

13) Haff, L. A. & Smirnov, I. P. (1997) Nucleic Acids Res. 25, 3749-3750.

14) Ross, P., Hall, L., Smirnov, I. P. & Haff, L. (1998) Nat. Biotechncl. 16, 1347-1351.

15) Fei Z, Ono T, Smith LM. (1998) MALDI-TOF mass spectrometric typing of single nucleotide polymorphisms with mass-tagged ddNTPs. *Nucleic Acids Res.* 26: 2827-2828.

16) Tang K, Fu DJ, Julien D, Braun A, Cantor CR, Koster H. (1999) Chip-based genotyping by mass spectrometry. *Proc. Natl. Acad. Sci. USA.* 96: 10016-10020.

17) Taranenko, N. I., Allman, S. L., Golovlev, V. V., Taranenko, N. V., Isola, N. R. & Chen, C. H. (1998) Nucleic Acids Res. 26, 2488-2490.

18) Ju J. Nucleic Acid Sequencing with Solid Phase Capturable Terminators. United States Patent No. 5,876,936, issued March 2, 1999.

19) Edwards, J. R., Itagaki, Y. & Ju, J. (2001) Nucleic Acids Res. 29, e104 (p1-5).

20) Tong, A. K. & Ju, J. (2002) Single nucleotide polymorphism detection by combinatorial fluorescence energy transfer tags and biotinylated dideoxynucleotides. Nucleic Acids Res. 30(5):e19.

21) Roskey MT, Juhasz P, Smirnov IP, Takach EJ, Martin SA, Haff LA. (1996) DNA sequencing by delayed extraction-matrix-assisted laser desorption/ ionization time of flight mass spectrometry. *Proc. Natl. Acad. Sci. USA.* 93: 4724-4729.

22) Hanson, E. H., Imperatore, G. & Burke, W. (2001) Am. J. Epidem. 154, 193-206.

23) Langer PR, Waldrop AA, Ward DC. (1981) Enzymatic synthesis of biotin-labeled polynucleotides: novel nucleic acid affinity probes. *Proc. Natl. Acad. Sci. USA.* 78: 6633-6637.

24) Hawkins, T. L., O'Connor-Morin, T., Roy, A. & Santillan, C. (1994) Nucleic Acids Res. 22, 4543-4544.

25) Uhlen, M. (1989) Nature, 340, 733-734.

26) Wahlberg, J., Lunderberg, J., Hultman, T. & Uhlen, M. (1990) Proc. Natl. Acad. Sci. USA. 87, 6569-6573.

27) Tong, X., Smith LM (1992) Solid-Phase Method for the Purification of DNA Sequencing Reactions. *Anal. Chem.* 64: 2672-2677.

28) Schneider K, Chait BT. (1995) Increased stability of nucleic acids containing 7-deaza-guanosine and 7-deaza-

adenosine may enable rapid DNA sequencing by matrix-assisted laser desorption mass spectrometry. *Nucleic Acids Res.* 23: 1570-1575.

29) Monforte JA, Becker CH (1997) High-throughput DNA analysis by time-of-flight mass spectrometry. *Nat Medicine.* 3(3): 360-362.

30) Hilton, G. C., Martinis, J. M., Wollman, D. A., Irwin, K. D., Dulcie, L. L., Gerber, D., Gillevet, P. M. & Twerenbold, D. (1998) Nature 391, 672-675.

31) Bergseid M, Baytan AR, Wiley JP, Ankener WM, Stolowitz, Hughs KA, Chestnut JD (Nov. 2000) Small-molecule base chemical affinity system for the purification of proteins. *BioTechniques* 29: 1126-1133.

32) Rosenblum BB, Lee LG, Spurgeon SL, Khan SH, Menchen SM, Heiner CR, Chen SM. (1997) New dye-labeled terminators for improved DNA sequencing patterns. *Nucleic Acids Res.* 25: 4500-4504.

33) Zhu Z, Chao J, Yu H, Waggoner AS. (1994) Directly labeled DNA probes using fluorescent nucleotides with different length linkers. *Nucleic Acids Res.* 22: 3418-3422.

34) Prober JM, Trainor GL, Dam RJ, Hobbs FW, Robertson CW, Zagursky RJ, Cocuzza AJ, Jensen MA, Baumeister K. (1987) A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* 238: 336-341.

35) Lee LG, Connell CR, Woo SL, Cheng RD, Mcardle BF, Fuller CW, Halloran ND, Wilson RK. (1992) DNA sequencing with dye-labeled terminators and T7 DNA -polymerase-effect of dyes and dNTPs on incorporation of dye-terminators and probability analysis of termination fragments. *Nucleic Acids Res.* 20: 2471-2483.

36) Hobbs FW Jr, Cocuzza AJ. Alkynylamino-Nucleotides. United States Patent No. 5,047,519, issued September 10, 1991.

37) Burgess K, Cook D. (2000) *Chemical Reviews.* 100: 2047-2060.

38) Olejnik J, Sonar S, Krzymanska-Olejnik E, Rothschild KJ. (1995) Photocleavable biotin derivatives: a versatile approach for the isolation of biomolecules. *Proc. Natl. Acad. Sci. USA.* 92: 7590-7594.

39) Olejnik J, Ludemann HC, Krzymanska-Olejnik E, Berkenkamp S, Hillenkamp F, Rothschild KJ. (1999) Photocleavable peptide-DNA conjugates: synthesis and applications to DNA analysis using MALDI-MS. *Nucleic Acids Res.* 27: 4626-4631.

40) Jurinke C, van de Boom D, Collazo V, Luchow A, Jacob A, Koster H. (1997) Recovery of nucleic acids from immobilized biotin-streptavidin complexes using ammonium hydroxide and applications in MALDI-TOF mass spectrometry. *Anal. Chem.* 69: 904-910.

41) Maudling DR, Lotts KD, Robinson SA. (1983) New procedure for making 2-(chloromethyl)-4-nitrotoluene. *J. Org. Chem.* 48: 2938.

42) Rolla F. (1982) Sodium-borohydride reactions under phase-transfer conditions - reduction of azides to amines. *J. Org. Chem.* 47: 4327-4329.

43) Woolley AT, Mathies RA. (1994) Ultra-high-speed DNA fragment separations using microfabricated capillary array electrophoresis chips. *Proc. Natl. Acad. Sci. USA.* 91: 11348-11352.

44) Collins FS, Green ED, Guttmacher, AE and Guyer MS. (2003) *Nature* 422: 835-847.

45) Jurinke C, van den Boom D, Cantor CR and Koster H. (2002) *Methods Mol. Biol.*187: 179-192

46) Kim S et al. (2003) *Analytical Biochemistry* 316: 251-258

What is claimed is:

1. A method for determining the identity of a nucleotide present at a predetermined site in a DNA whose sequence immediately 3′ of such predetermined site is known which comprises:

(a) treating the DNA with an oligonucleotide primer whose sequence is complementary to such known sequence so that the oligonucleotide primer hybridizes to the DNA and forms a complex in which the 3′ end of the oligonucleotide primer is located immediately adjacent to the predetermined site in the DNA;

(b) simultaneously contacting the complex from step (a) with four different labeled dideoxynucleotides, in the presence of a polymerase under conditions permitting a labeled dideoxynucleotide to be added to the 3′ end of the primer so as to generate a labeled single base extended primer, wherein each of the four different labeled dideoxynucleotides (i) is complementary to one of the four nucleotides present in the DNA and (ii) has a molecular weight which can be distinguished from the molecular weight of the other three labeled dideoxynucleotides using mass spectrometry; and

(c) determining the difference in molecular weight between the labeled single base extended primer and the oligonucleotide primer so as to identify the dideoxynucleotide incorporated into the single base extended primer and thereby determine the identity of the

nucleotide present at the predetermined site in the DNA.

2. The method of claim 1, wherein each of the four labeled dideoxynucleotides comprises a chemical moiety attached to the dideoxynucleotide by a different linker which has a molecular weight different from that of each other linker.

3. The method of claim 1 which further comprises after step (b) the steps of:

(i) contacting the labeled single base extended primer with a surface coated with a compound that specifically interacts with a chemical moiety attached to the dideoxynucleotide by a linker so as to thereby capture the extended primer on the surface; and

(ii) treating the labeled single base extended primer so as to release it from the surface.

4. The method of claim 3 which further comprises after step (i) the step of treating the surface to remove primers that have not been extended by a labeled dideoxynucleotide.

5. The method of claim 1, wherein step (c) comprises determining the difference in mass between the labeled single base extended primer and an internal mass calibration standard added to the extended primer.

6. The method of claim 3, wherein the interaction between the chemical moiety attached to the

dideoxynucleotide by the linker and the compound on the surface comprises a biotin-streptavidin interaction, a phenylboronic acid-salicylhydroxamic acid interaction, or an antigen-antibody interaction.

7. The method of claim 3, wherein the step of releasing the labeled single base extended primer from the surface comprises disrupting the interaction between the chemical moiety attached to the dideoxynucleotide by the linker and the compound on the surface.

8. The method of claim 7, wherein the interaction is disrupted by a means selected from the group consisting of one or more of a physical means, a chemical means, a physical chemical means, heat, and light.

9. The method of claim 2, wherein the linker is attached to the dideoxynucleotide at the 5-position of cytosine or thymine or at the 7-position of adenine or guanine.

10. The method of claim 3, wherein the step of releasing the labeled single base extended primer from the surface comprises cleaving the linker between the chemical moiety and the dideoxynucleotide.

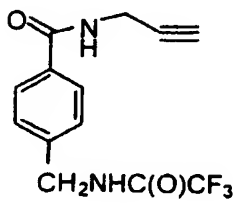
11. The method of claim 10, where the linker is cleaved by a means selected from the group consisting of one or more of a physical means, a chemical means, a physical chemical means, heat, and light.

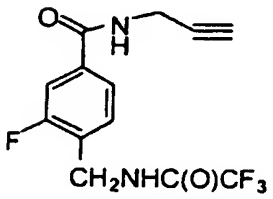
12. The method of claim 11, wherein the linker is cleaved by light.

13. The method of claim 2, wherein the linker comprises a derivative of 4-aminomethyl benzoic acid, a 2-nitrobenzyl group, or a derivative of a 2-nitrobenzyl group.

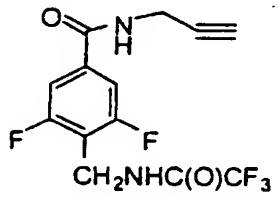
14. The method of claim 13, wherein the linker comprises one or more fluorine atoms.

15. The method of claim 14, wherein the linker is selected from the group consisting of:





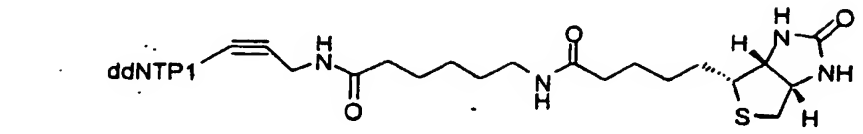
and

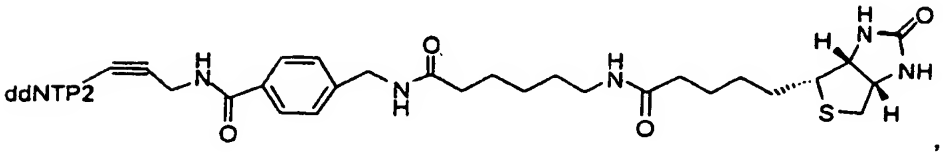


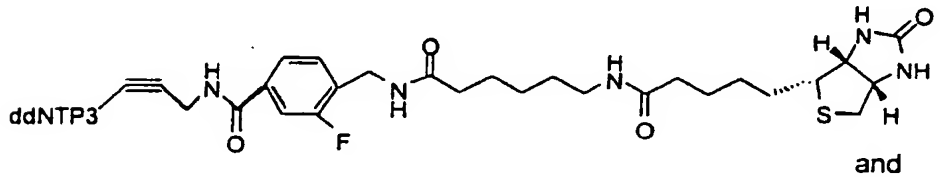
16. The method of claim 3, wherein the chemical moiety comprises biotin, the labeled dideoxynucleotide is a biotinylated dideoxynucleotide, the labeled single base extended primer is a biotinylated single base extended primer, and the surface is a streptavidin-coated solid surface.

17. The method of claim 16, wherein the biotinylated dideoxynucleotide is selected from the group consisting of ddATP-11-biotin, ddCTP-11-biotin, ddGTP-11-biotin, and ddTTP-16-biotin.

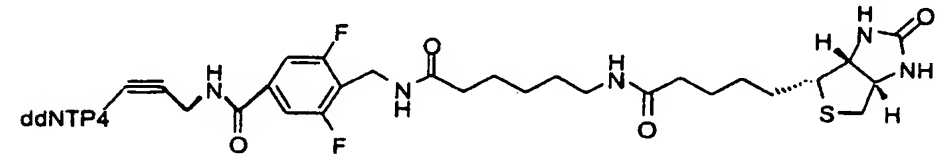
18. The method of claim 16, wherein the biotinylated dideoxynucleotide is selected from the group consisting of:





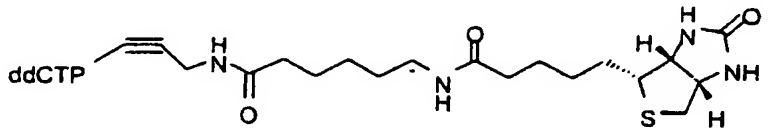


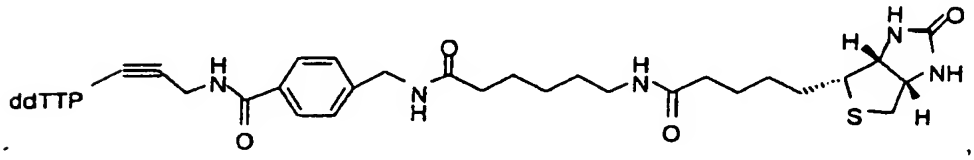
and

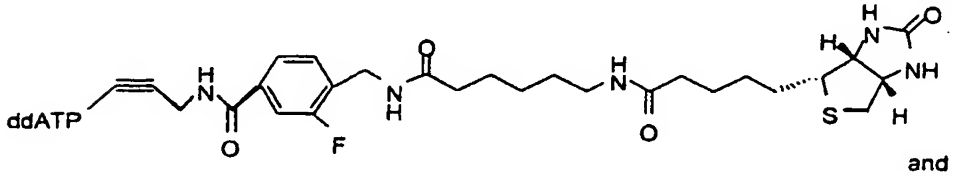


wherein ddNTP1, ddNTP2, ddNTP3, and ddNTP4 represent four different dideoxynucleotides.

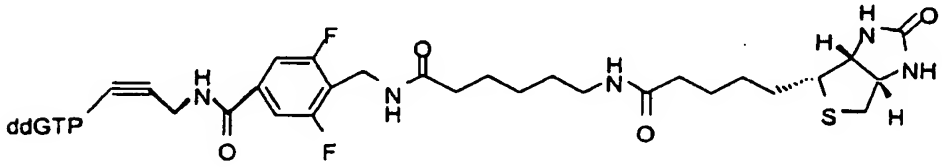
19. The method of claim 18, wherein the biotinylated dideoxynucleotide is selected from the group consisting of:



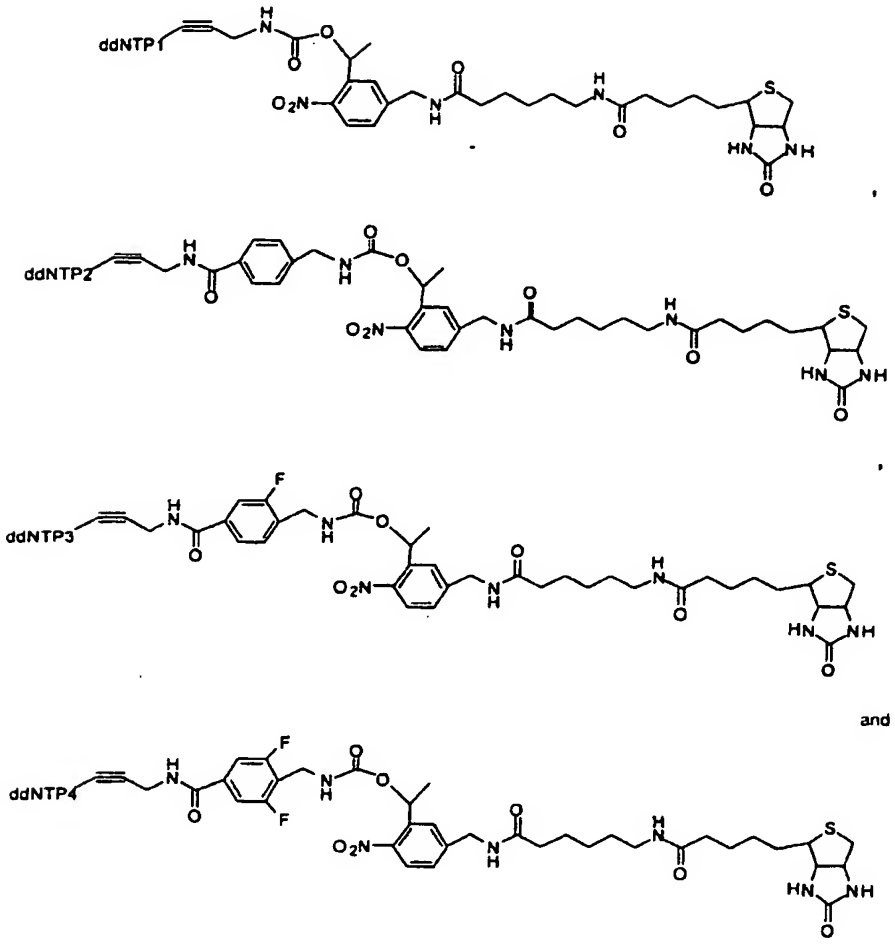




and

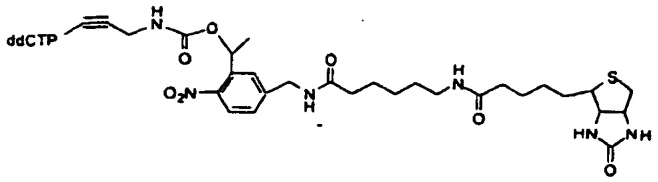


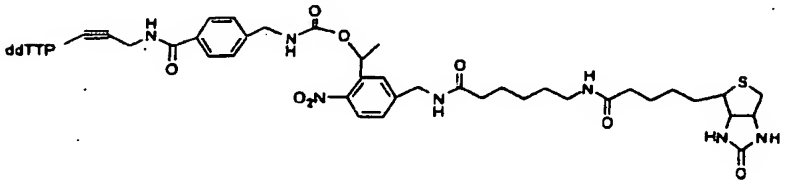
20. The method of claim 16, wherein the biotinylated dideoxynucleotide is selected from the group consisting of:

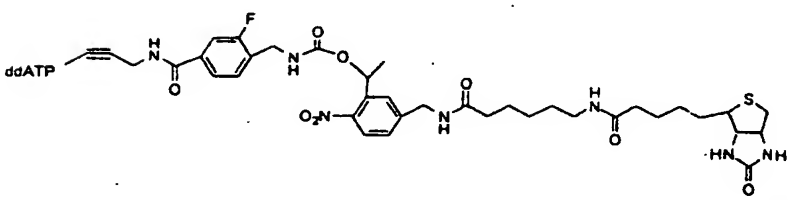


wherein ddNTP1, ddNTP2, ddNTP3, and ddNTP4 represent four different dideoxynucleotides.

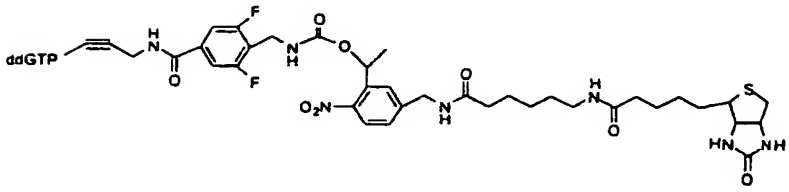
21. The method of claim 20, wherein the biotinylated dideoxynucleotide is selected from the group consisting of:







and



5

22. The method of claim 16, wherein the streptavidin-coated solid surface is a streptavidin-coated magnetic bead or a streptavidin-coated silica glass.

23. The method of claim 1, wherein steps (a) and (b) are performed in a single container or in a plurality of connected containers.

24. A method for determining the identity of nucleotides present at a plurality of predetermined sites, which comprises carrying out the method of claim 3 using a plurality of different primers each having a molecular weight different from that of each other primer, wherein a different primer hybridizes adjacent to a different predetermined site.

25. The method of claim 24, wherein different linkers each having a molecular weight different from that of each other linker are attached to the different dideoxynucleotides to increase mass separation between different labeled single base extended primers and thereby increase mass spectrometry resolution.

FIGURE 1

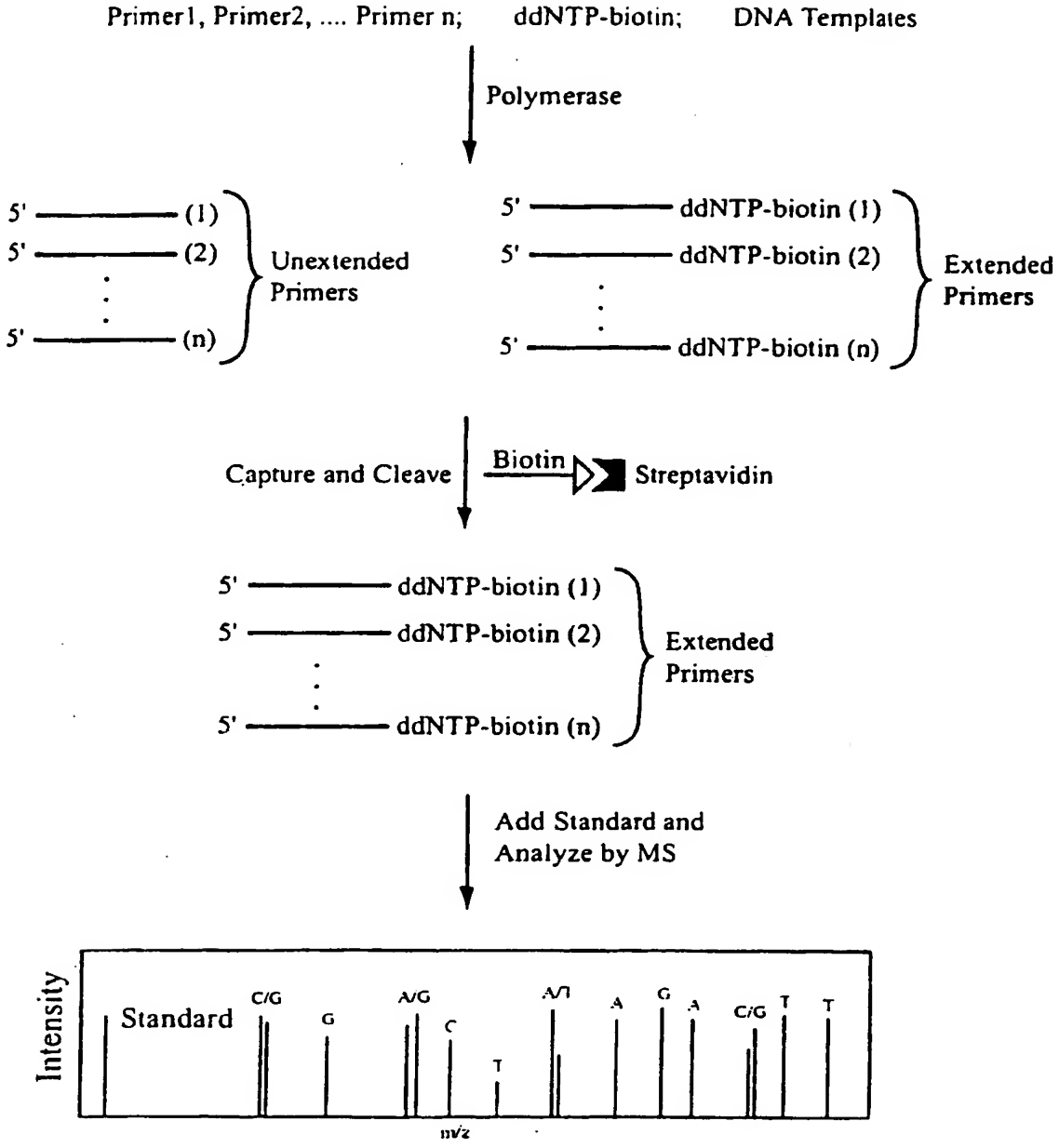

Primer1, Primer2, ... Primer n;
ddNTP-biotin;
DNA Templates
Polymerase
5' (1)
5' (2)
5' (n)
Unextended Primers
5' ddNTP-biotin (1)
5' ddNTP-biotin (2)
5' ddNTP-biotin (n)
Extended Primers
Capture and Cleave
Biotin
Streptavidin
5' ddNTP-biotin (1)
5' ddNTP-biotin (2)
5' ddNTP-biotin (n)
Extended Primers
Add Standard and Analyze by MS
Intensity
Standard
C/G
G
A/G
C
T
A/T
A
G
A
C/G
T
T
m/z

FIGURE 2A

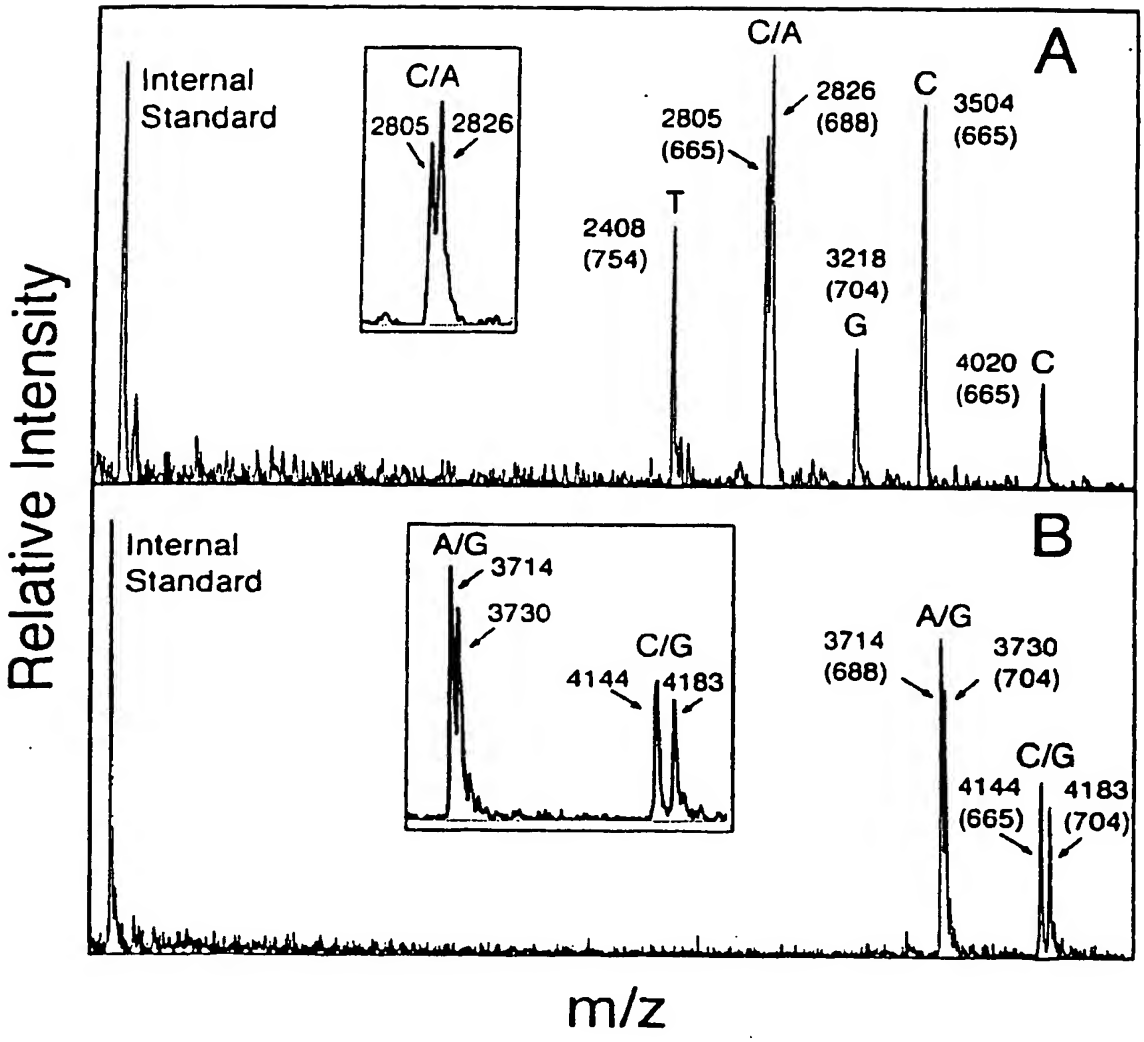


FIGURE 2B

FIGURE 3

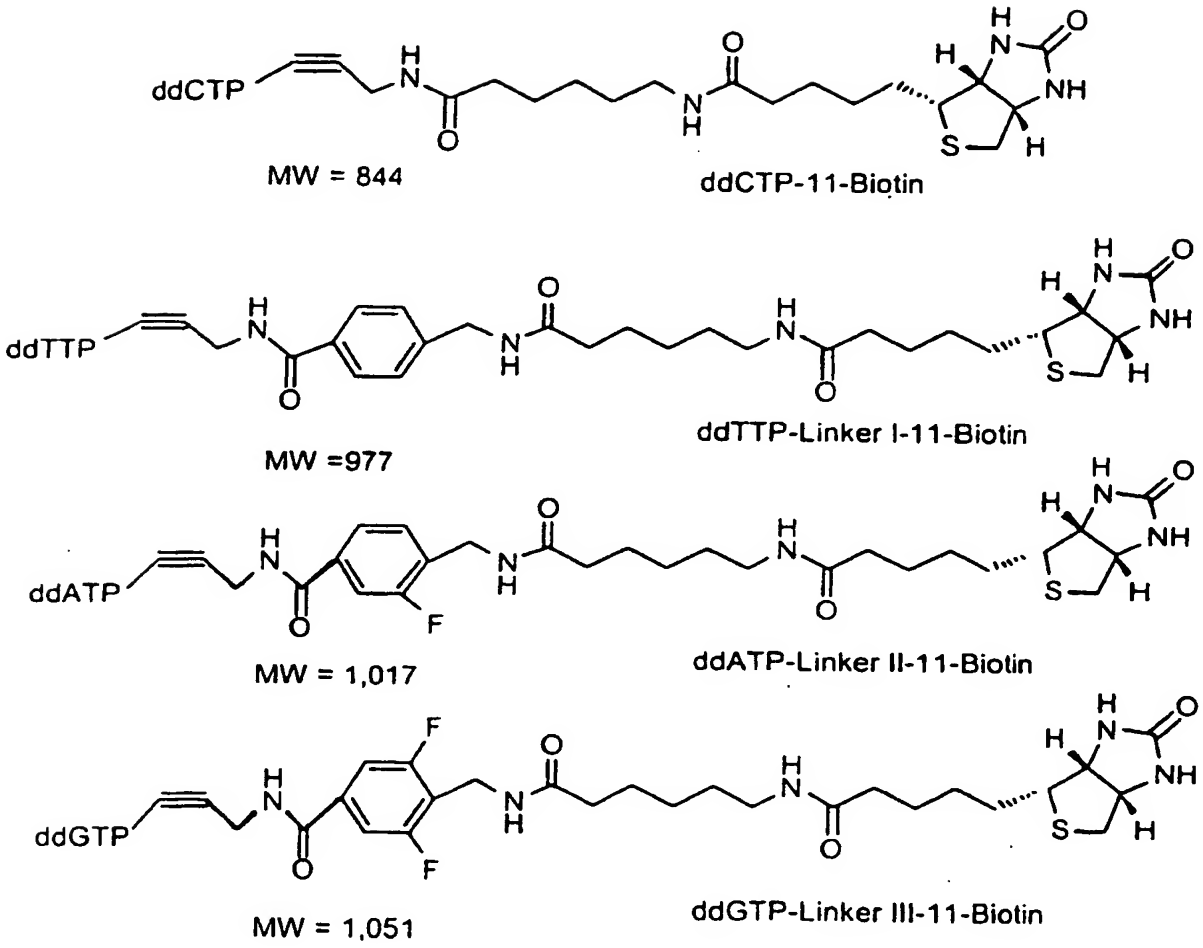
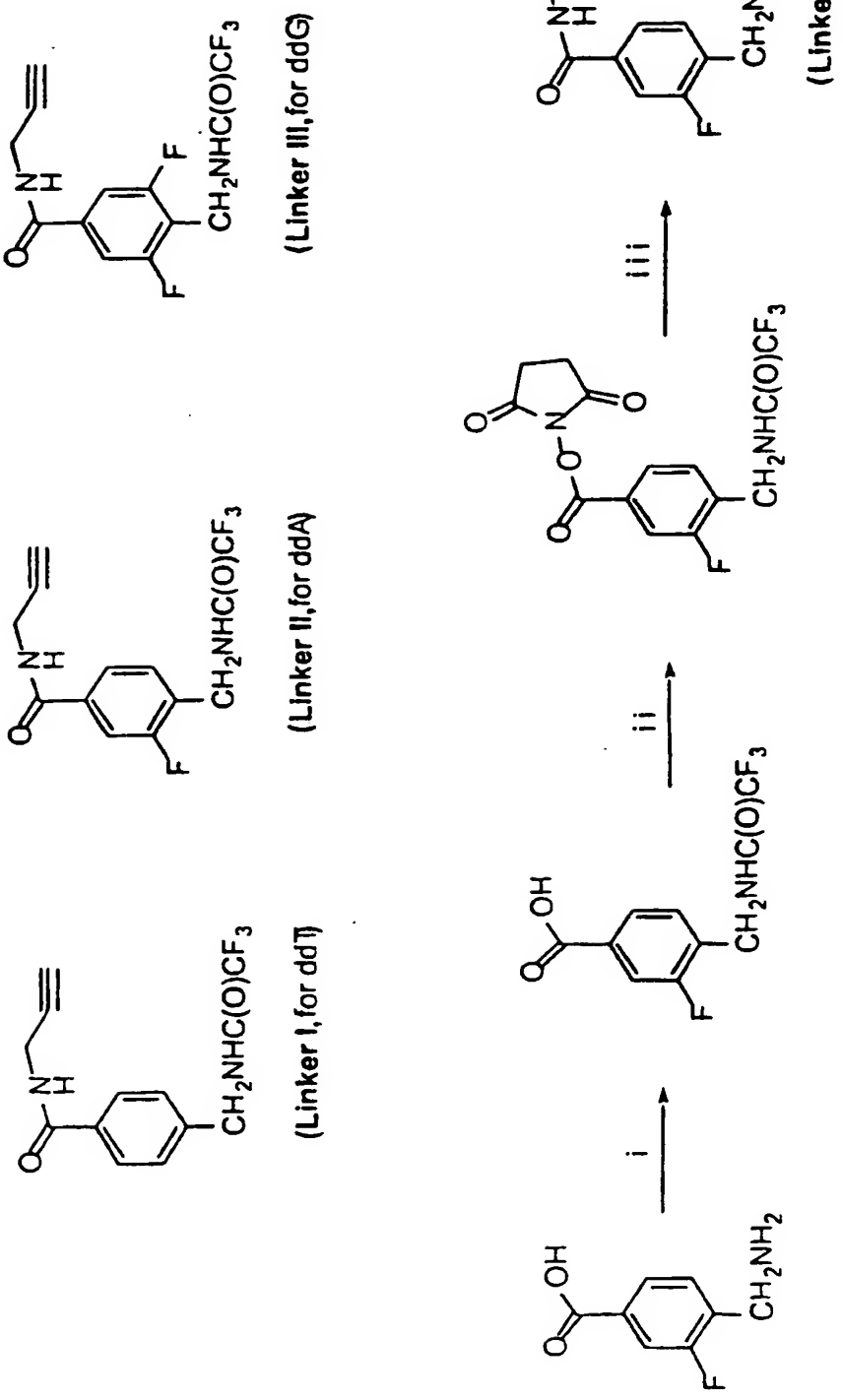

ddCTP
MW = 844
ddCTP-11-Biotin
ddTTP
MW =977
ddTTP-Linker I-11-Biotin
ddATP
F
MW = 1,017
ddATP-Linker II-11-Biotin
ddGTP
F
F
MW = 1,051
ddGTP-Linker III-11-Biotin

FIGURE 4


i
ii
iii
CH_2NH_2
$CH_2NHC(O)CF_3$
(Linker I, for ddT)
(Linker II, for ddA)
(Linker III, for ddG)
(Linker II, for ddA)

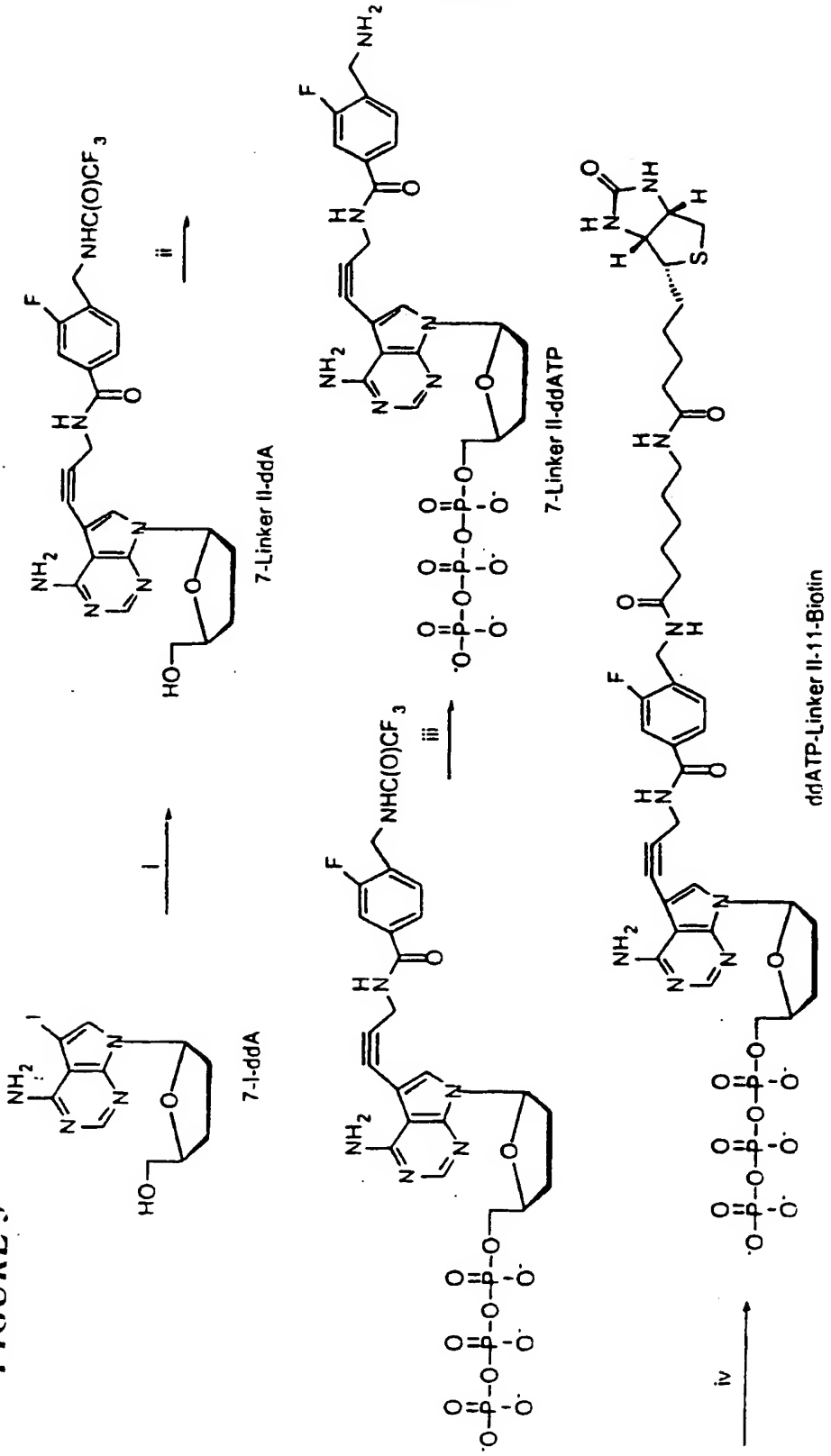

7-I-ddA
7-Linker II-ddA
7-Linker II-ddATP
ddATP-Linker II-11-Biotin
i
ii
iii
iv

FIGURE 5

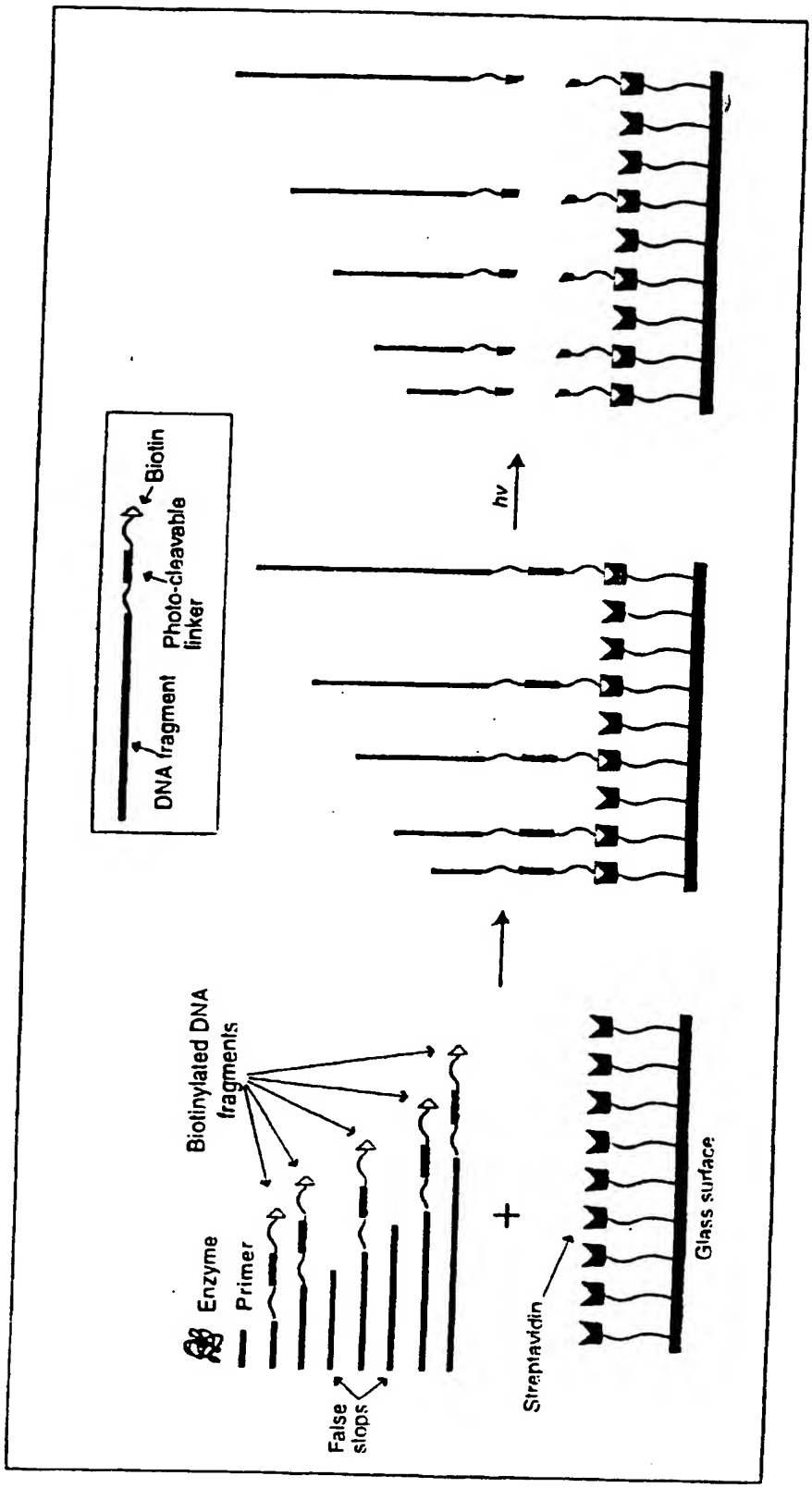

Biotin
Photo-cleavable linker
DNA fragment
hν
Biotinylated DNA fragments
Enzyme
Primer
False stops
Streptavidin
Glass surface

FIGURE 6

FIGURE 7

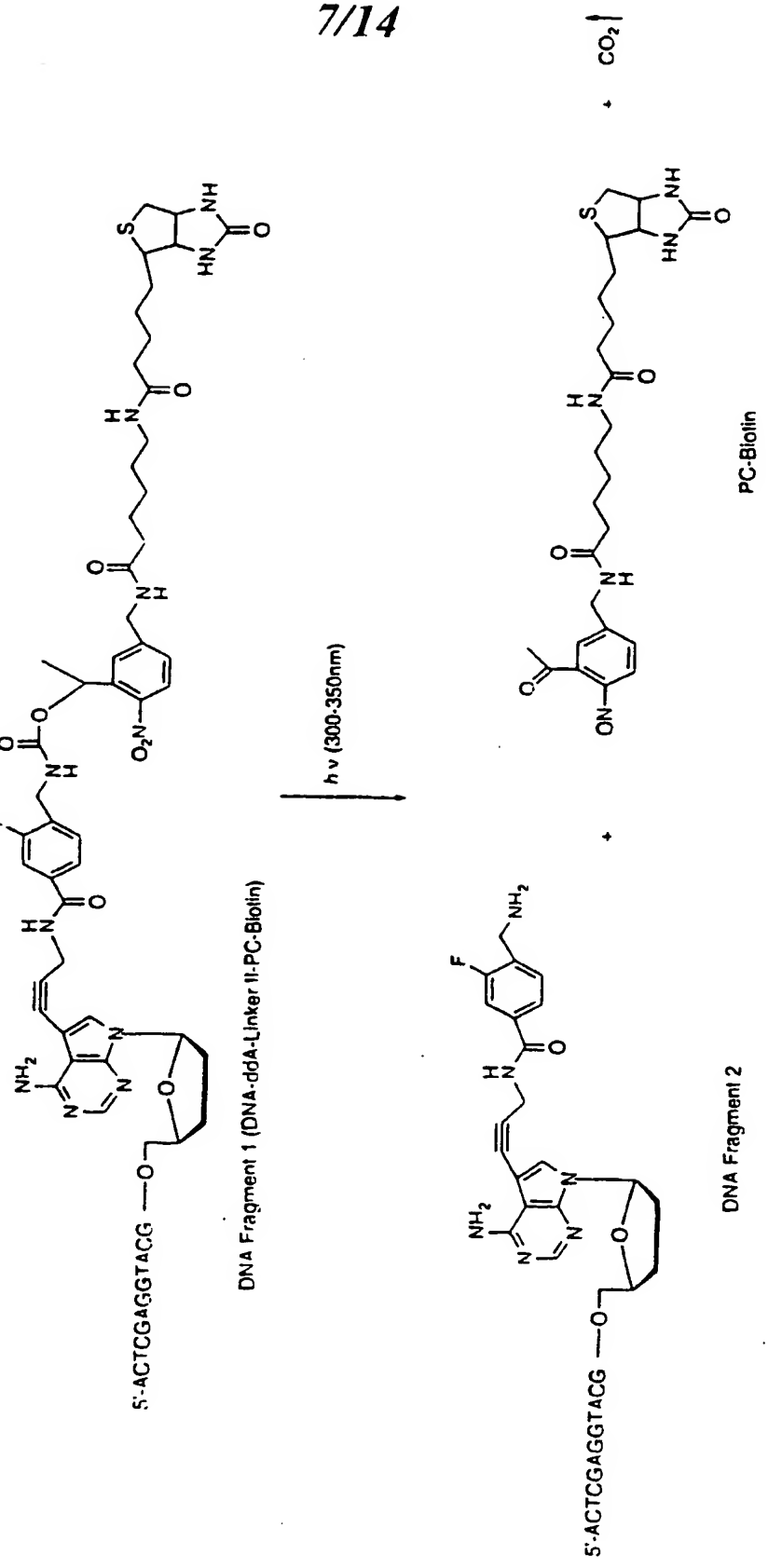

5'-ACTCGAGGTACG
DNA Fragment 1 (DNA-ddA-Linker II-PC-Biotin)
hν (300-350nm)
PC-Biotin
CO2
5'-ACTCGAGGTACG
DNA Fragment 2

FIGURE 8

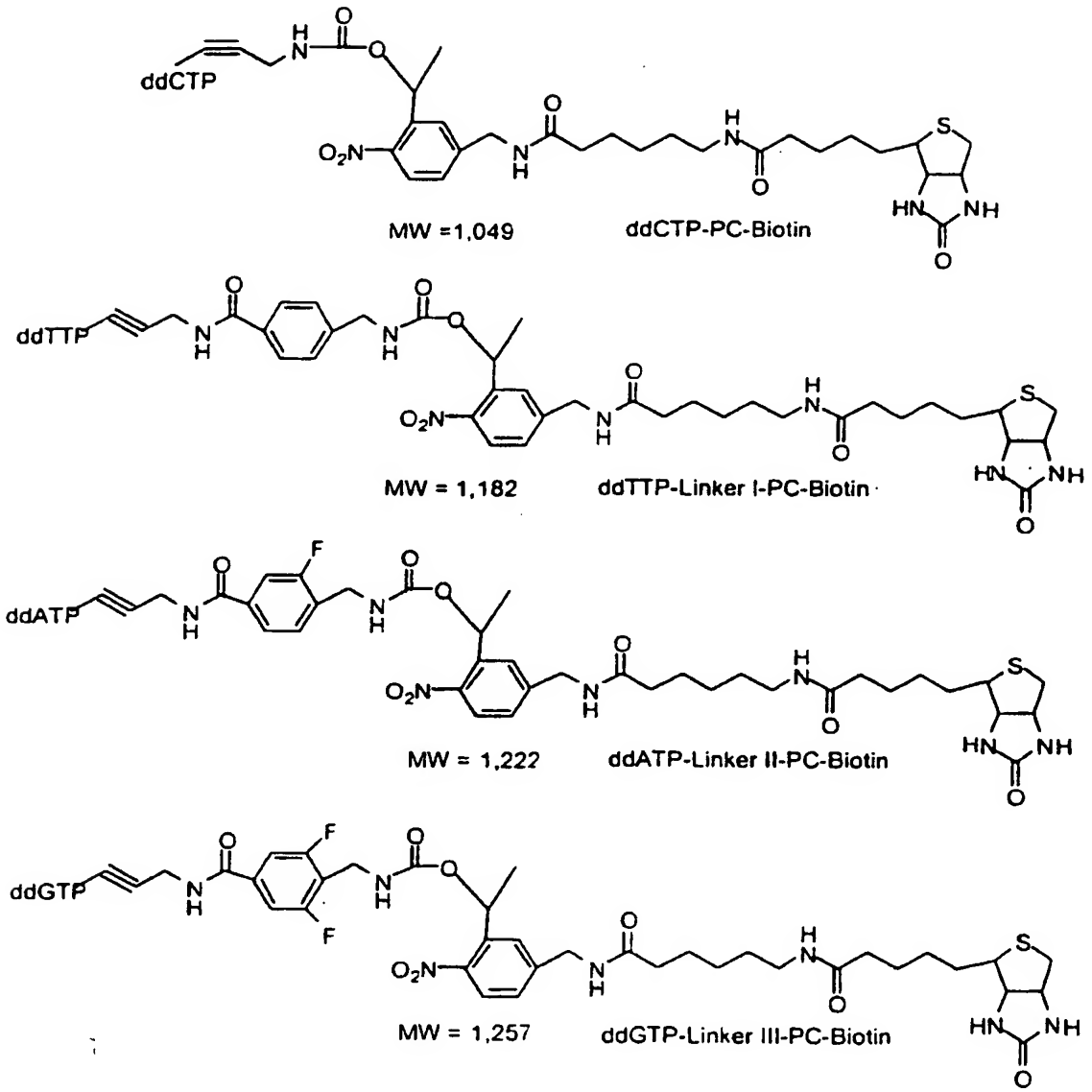
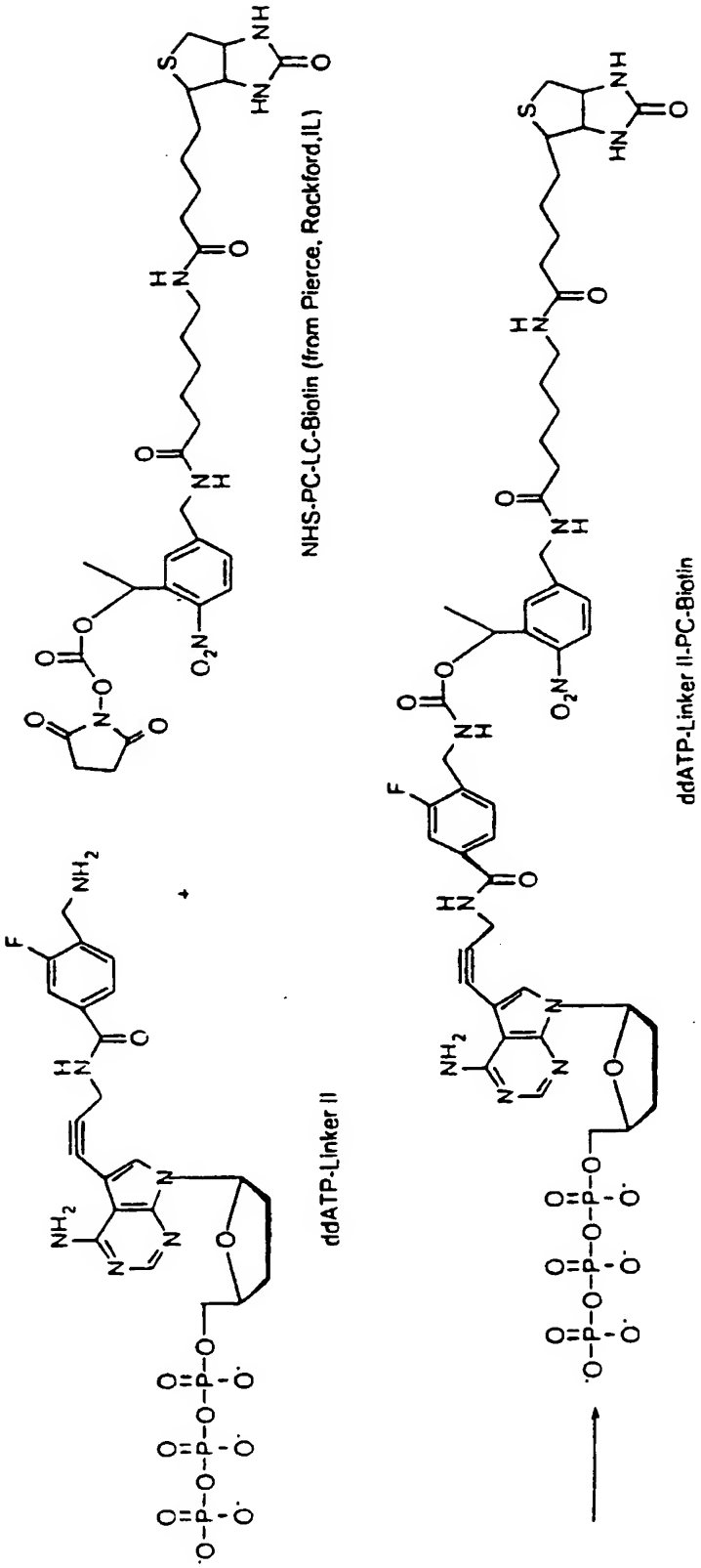

ddCTP
MW =1,049
ddCTP-PC-Biotin
ddTTP
MW = 1,182
ddTTP-Linker I-PC-Biotin
ddATP
MW = 1,222
ddATP-Linker II-PC-Biotin
ddGTP
MW = 1,257
ddGTP-Linker III-PC-Biotin

FIGURE 9


NHS-PC-LC-Biotin (from Pierce, Rockford,IL)
ddATP-Linker II
ddATP-Linker II-PC-Biotin

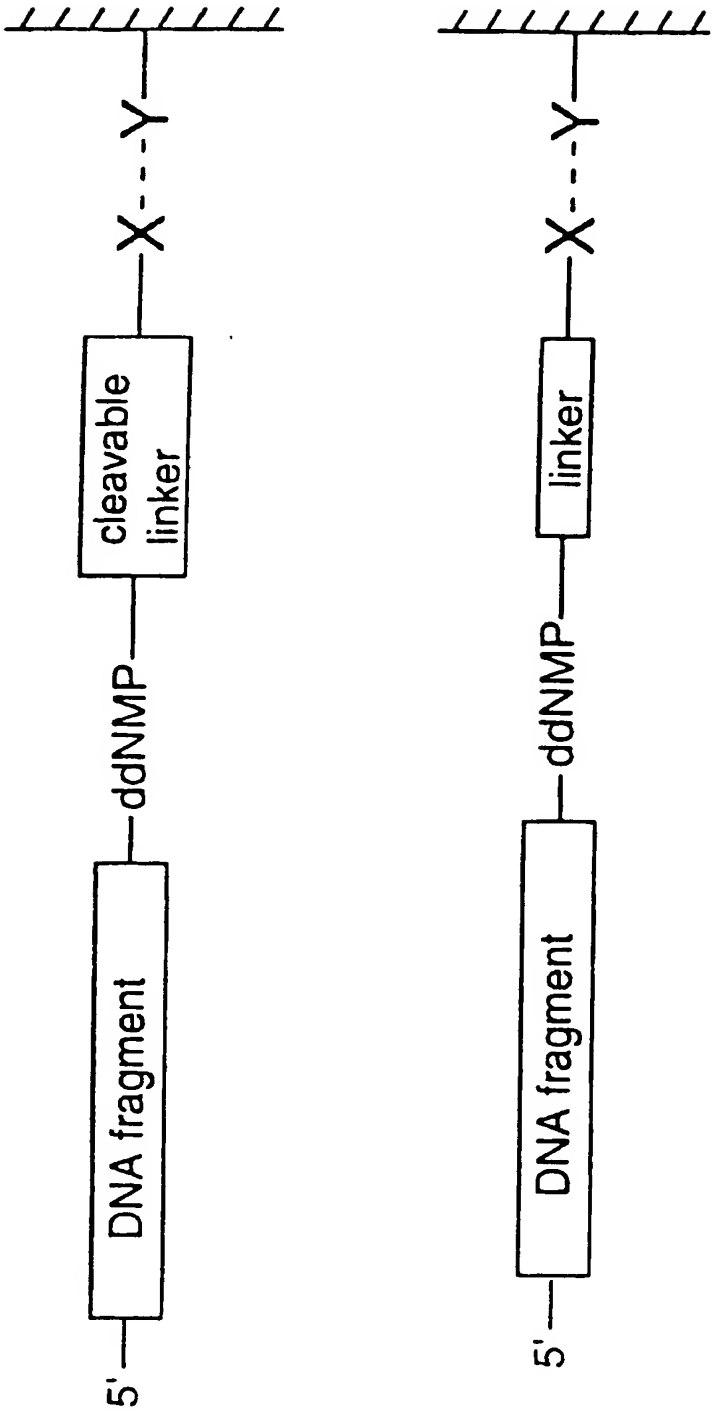

X
Y
cleavable linker
ddNMP
DNA fragment
5'
X
Y
linker
ddNMP
DNA fragment
5'

FIGURE 10

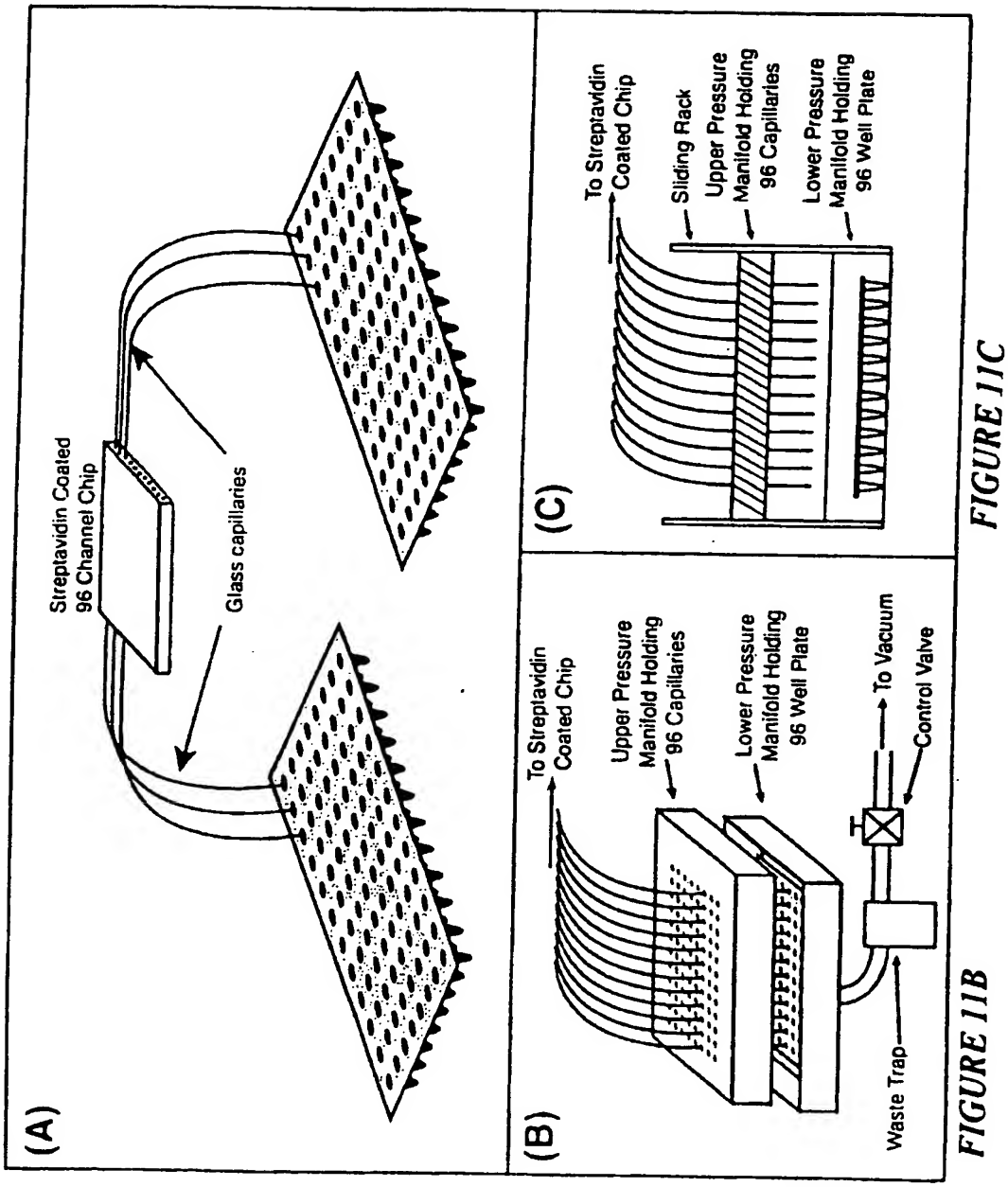


FIGURE 11A

FIGURE 11B

FIGURE 11C

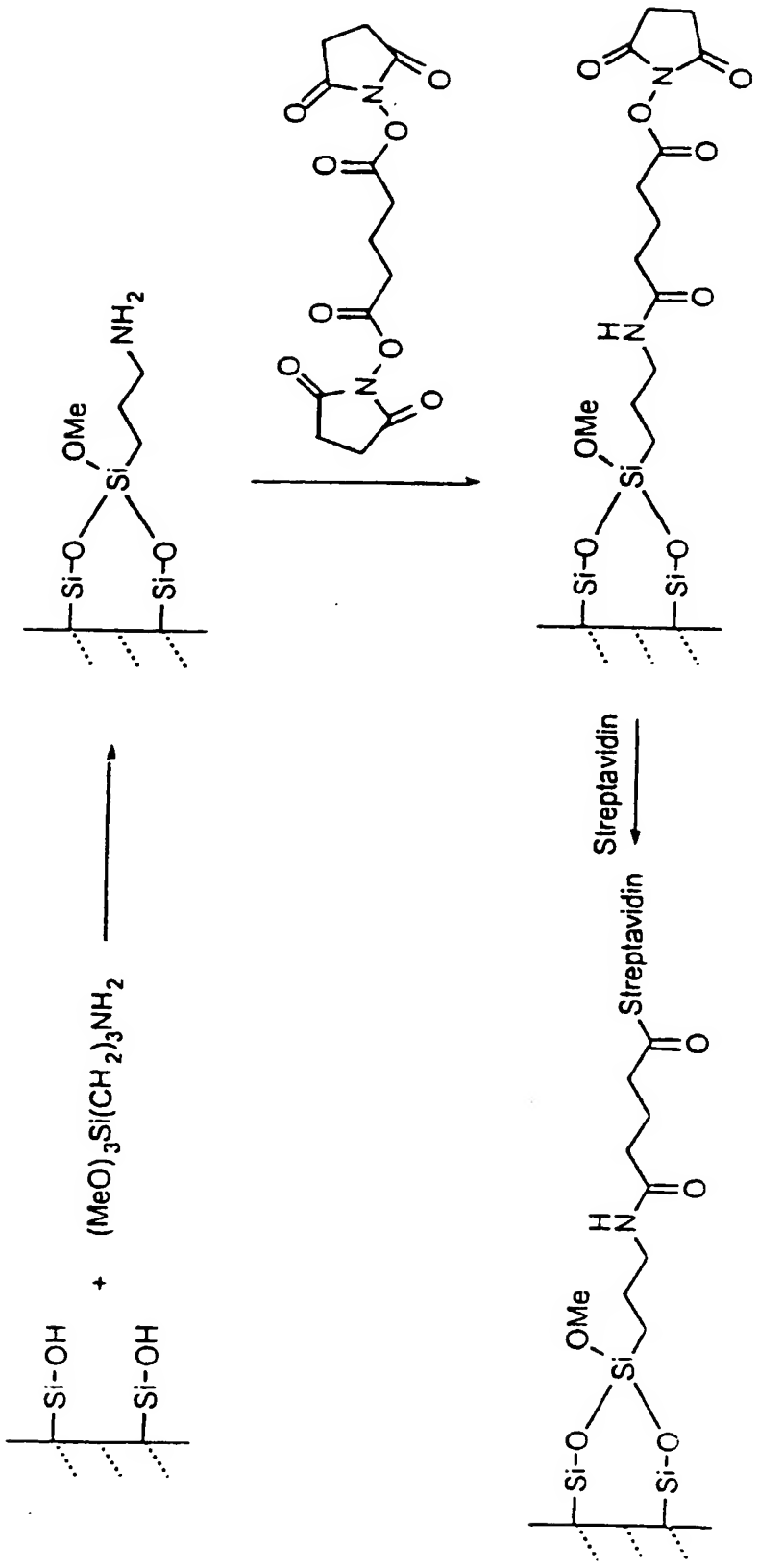

Si-OH
Si-OH
+
(MeO)3Si(CH2)3NH2
OMe
Si
Si-O
Si-O
NH2
Streptavidin
Streptavidin
H
N
O

FIGURE 12

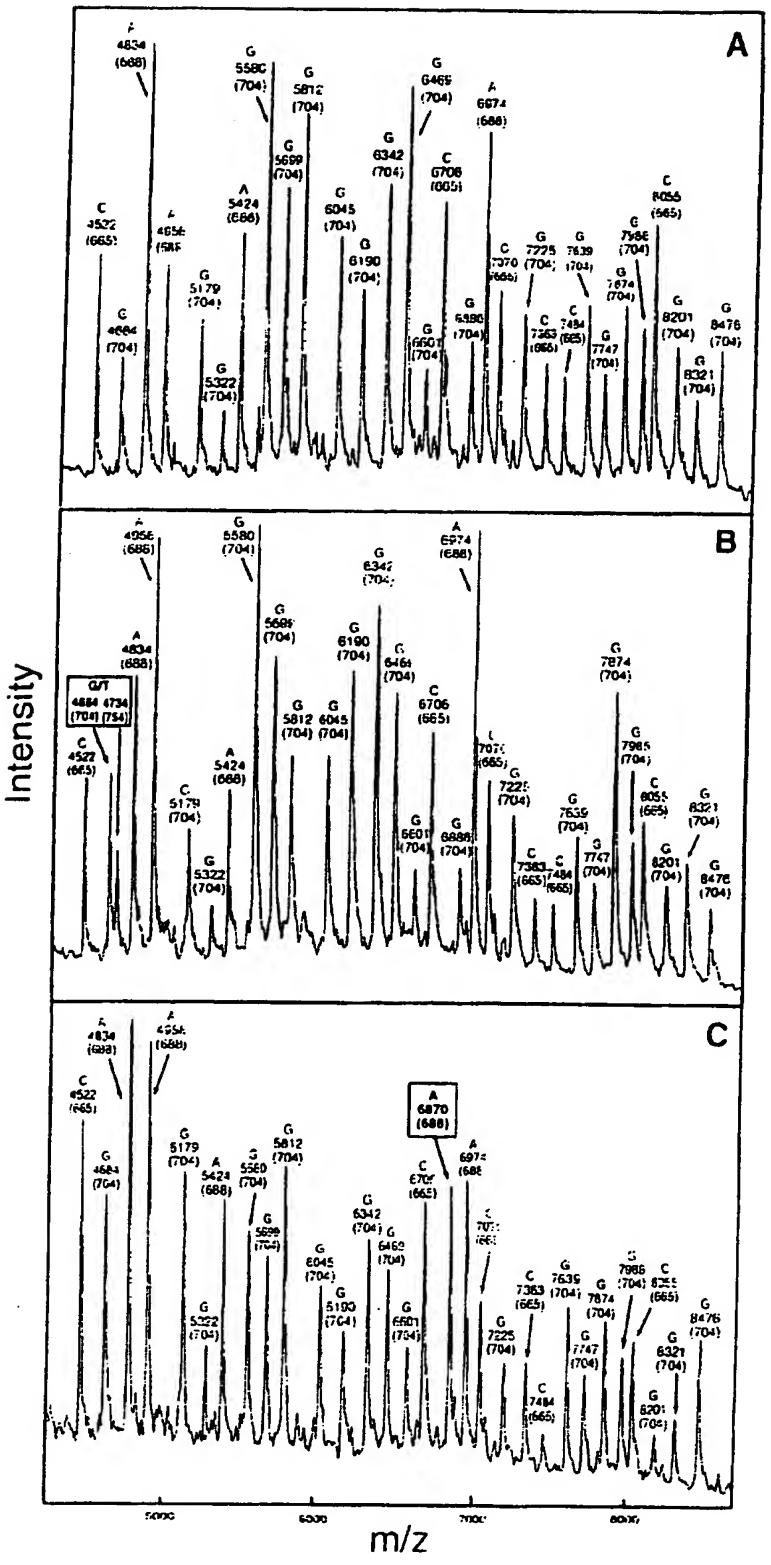


FIGURE 13A

FIGURE 13B

FIGURE 13C

FIGURE 14A

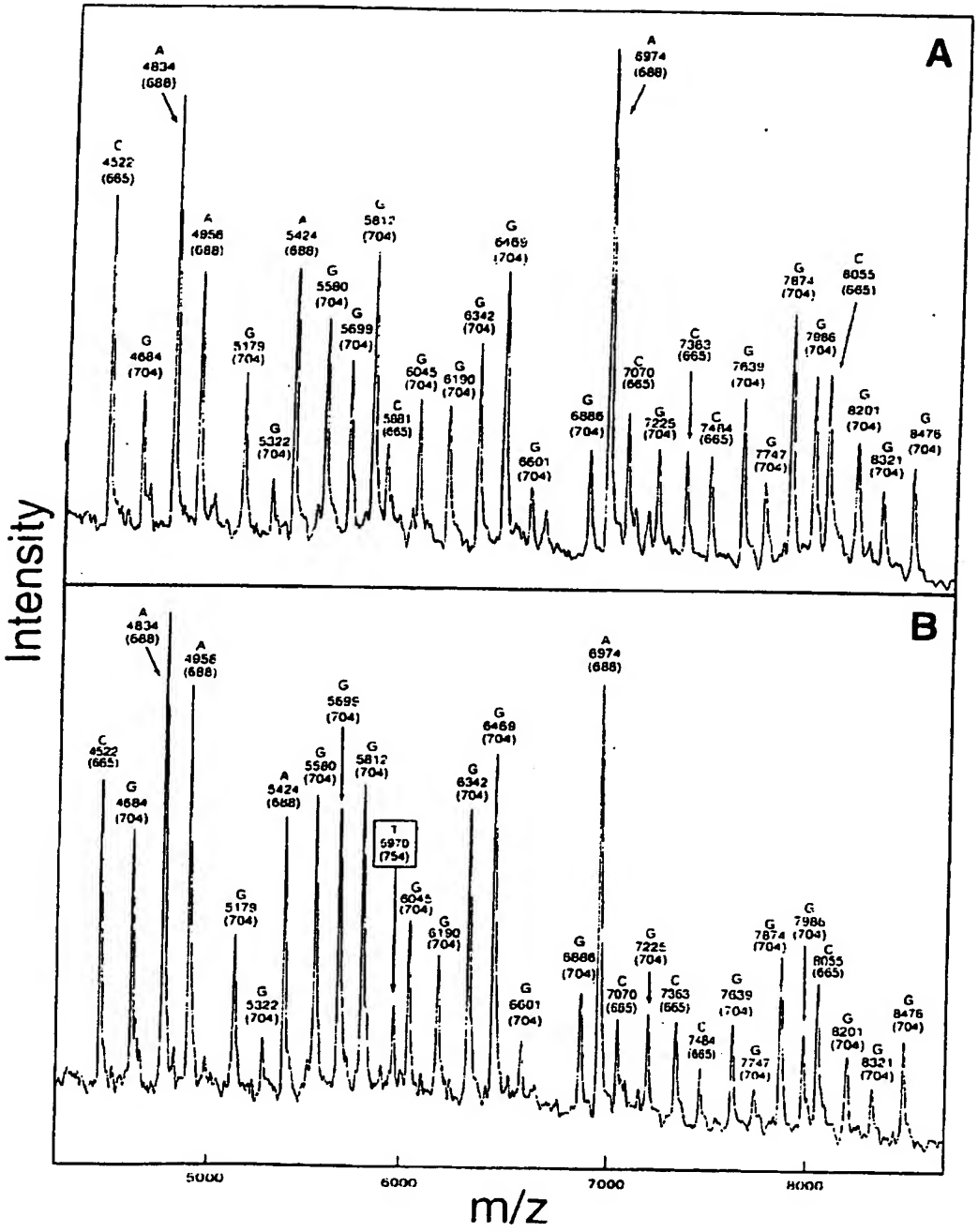


FIGURE 14B

SEQUENCE LISTING

<110> The Trustees of Columbia University in the City of New York et al.

<120> Multiplex Genotyping Using Solid Phase Capturable Dideoxynucleotides And Mass Spectrometry

<130> 0575/66833-A-PCT

<140> NOT YET KNOWN
<141> HEREWITH

<160> 14

<170> PatentIn version 3.2

<210> 1
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 1
ctacccccag aacatcacc 19

<210> 2
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 2
gcactacctc ttcatgggtg cc 22

<210> 3
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 3
catcagtcac atacccca 18

<210> 4
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 4
cagtgaacat gtgatcccac cc 22

<210> 5
<211> 13
<212> DNA
<213> Artificial Sequence

<220>
<223> internal mass standard

<400> 5
ttttttcttt tct 13

<210> 6
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 6
ggggaagagc agagatatac gt 22

<210> 7
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 7
ggggctccac acggcgactc tcat 24

<210> 8
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 8
agaggatcca accgagac 18

<210> 9
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 9
tggtggtagg tgatgttgat gta 23

<210> 10
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 10
cacattgtca aggacgtacc cg 22

<210> 11
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 11
tacccgccgt acttggcctc 20

<210> 12
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> primer

<400> 12
tccacgcaca aacacggaca g 21

<210> 13
<211> 100
<212> DNA
<213> Artificial Sequence

<220>
<223> template

<400> 13
taccckgagg ccaagtacgg cgggtacgtc cttgacaatg tgtacatcaa catcacctac 60

caccatgtca gtctcggttg gatcctctat tgtgtccggg 100

<210> 14
<211> 110
<212> DNA
<213> Artificial Sequence

<220>
<223> template

<400> 14
gaaggagaca cgcggccaga gagggtcctg tccgtgtttg tgcgtggagt ttcgacaagg 60

cagggtcatc taatggtgat gagtcctatc cttttctctt cgttctccgt 110

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US03/21818

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12Q 1/68, C12P 19/34, C07H 21/02, 21/04, 19/00
US CL : 435/9, 7.1,91.1,91.2; 536 22.1,23.1 24.3,24.31,24.32,24.33

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/9, 7.1,91.1,91.2; 536 22.1,23.1 24.3,24.31,24.32,24.33

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 5,849,542 A (REEVE et al) 15 December 1998 (15.12. 1998), see entire document esp. abstract.	1-25
Y	US 5,885,775 A (HAFF et al) 23 March 1999 (23.03. 1999), see entire document esp. abstract.	1-25

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
08 September 2003 (08.09.2003)

Date of mailing of the international search report
26 SEP 2003

Name and mailing address of the ISA/US
Mail Stop PCT, Attn: ISA/US
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
Facsimile No. (703)305-3230

Authorized officer
Jeffrey Siew
Janice Ford
for
Telephone No. 703-308-0196

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

PCT/US03/21818

Continuation of B. FIELDS SEARCHED Item 3:
EAST -ALL DBS, STN-BIOSIS, MEDLINE, CANCERLIT, BIOTECHDS, LIFESCI, CAPLUS, EMBASE
search terms: mass spectrometer, primer, ddntp